UCATIONAL and YCHOLOGICAL MEASUREMENT

VOLUME THREE, NUMBER THREE, AUTUMN

An Experimental Analysis of Five Tests of "Appreciation of Humor:" H. J. EYSENCK 191

The Development and Evaluation of a Measure of Counseling Interview Procedure: E. H. PORTER, JR. 215

A Comparison of Interest Measurement by the Kuder Preference Record and the Strong Vocational Interest Blanks for Men and Women: J. R. WITTENBORN, FRANCES ORALIND TRIGGS and DANIEL D. FEDER 239

An Evaluation of Certain Tests in Predicting Mechanic Learner Achievement: ELDON E. JACOBSEN 259

The Use of Civil Service Examination Bulletins as a Counseling Aid: ARTHUR BURTON 269

Measurement Abstracts 273

News Notes 283

Recommendations of the Council of Guidance and Personnel Associations 287

Copyright, 1943, by
SCIENCE RESEARCH ASSOCIATES

STATEMENT OF THE OWNERSHIP, MANAGEMENT, CIRCULATION, ETC., REQUIRED BY THE ACTS OF CONGRESS OF AUGUST 24, 1912, AND MARCH 3, 1933, of EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT, published quarterly at Chicago, Illinois, for October 1, 1943. STATE OF Illinois, COUNTY OF Cook.

Before me, a Notary Public in and for the State and county aforesaid, personally appeared Walter A. Symons, who, having been duly sworn according to law, deposes and says that he is the Business Manager of the EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT, and that the following is, to the best of his knowledge and belief, a true statement of the ownership, management (and if a daily paper, the circulation), etc., of the aforesaid publication for the date shown in the above caption, required by the Act of August 24, 1912, as amended by the Act of March 3, 1933, embodied in section 537, Postal Laws and Regulations, printed on the reverse of this form, to wit:

1. That the names and addresses of the publisher, editor, managing editor, and business managers are: Publisher, SCIENCE RESEARCH ASSOCIATES, 1700 Prairie Avenue, Chicago. Editor, G. Frederic Kuder, 1700 Prairie Avenue, Chicago. Managing Editor, John R. Yale, 1700 Prairie Avenue, Chicago. Business Manager, Walter A. Symons, 1700 Prairie Avenue, Chicago.

2. That the owner is: (If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding one per cent or more of total amount of stock. If not owned by a corporation, the names and addresses of the individual owners must be given. If owned by a firm, company, or other unincorporated concern, its name and address, as well as those of each individual member, must be given.) SCIENCE RESEARCH ASSOCIATES, INC., 1700 Prairie Avenue, Chicago, Illinois. Ralph A. Bard, 208 S. LaSalle Street, Chicago, Illinois. Charles S. Boyd, Appleton Coated Paper Co., Appleton, Wisconsin. R. W. Glasner, 6499 W. 65th Street, Chicago, Illinois. Alfred E. Hamill, 208 S. LaSalle Street, Chicago, Illinois. Robert C. McNamara, 623 S. Wabash Avenue, Chicago, Illinois. John I. Shaw, 135 S. LaSalle Street, Chicago, Illinois. Lyle M. Spencer, 1700 Prairie Avenue, Chicago, Illinois. Mrs. Dorothy Bard, c/o Roy E. Bard, 134 S. LaSalle Street, Chicago, Illinois. Roy E. Bard, 134 S. LaSalle Street, Chicago, Illinois. George M. Bard II, c/o Ralph A. Bard, 208 S. LaSalle Street, Chicago, Illinois. Miss Janet Bard, c/o Ralph A. Bard, 208 S. LaSalle Street, Chicago, Illinois. Robert K. Burns, 1700 Prairie Avenue, Chicago, Illinois. Miss Grace M. Wagner, c/o Richard Wagner, 135 S. LaSalle Street, Chicago, Illinois. W. C. Winkel, c/o Modine Mfg. Co., Racine, Wisconsin. Louis Eisman, 4468 Lake Park Avenue, Chicago, Illinois. John R. Yale, 1700 Prairie Avenue, Chicago, Illinois.

3. That the known bondholders, mortgagees, and other security holders owning or holding 1 per cent or more of total amount of bonds, mortgages, or other securities are: None.

4. That the two paragraphs next above, giving the names of the owners, stockholders, and security holders, if any, contain not only the list of stockholders and security holders as they appear upon the books of the company but also, in cases where the stockholder or security holder appears upon the books of the company as trustee or in any other fiduciary relation, the name of the person or corporation for whom such trustee is acting, is given; also that the said two paragraphs contain statements embracing affiant's full knowledge and belief as to the circumstances and conditions under which stockholders and security holders who do not appear upon the books of the company as trustees, hold stock and securities in a capacity other than that of a bona fide owner; and this affiant has no reason to believe that any other person, association, or corporation has any interest direct or indirect in the said stock, bonds, or other securities than as so stated by him.

5. That the average number of copies of each issue of this publication sold or distributed, through the mails or otherwise, to paid subscribers during the twelve months preceding the date shown above is (not a daily publication). (This information is required from daily publications only.)

Signed: Walter A. Symons, business manager. Sworn to and subscribed before me this 10th day of September, 1943.—Dorotha Mocher.
(Seal) (My commission expires June 12, 1945.)

PRINTED IN THE UNITED STATES OF AMERICA
THE SCIENCE PRESS PRINTING COMPANY
LANCASTER, PENNSYLVANIA

AN EXPERIMENTAL ANALYSIS OF FIVE TESTS OF "APPRECIATION OF HUMOR"

H. J. EYSENCK
Psychology Department, Mill Hill Emergency Hospital
London

Introduction

IT IS generally agreed, among psychologists as well as among laymen, that "sense of humor" is an important and valuable personality trait. It has been equated with "insight" (2), and it has been made into a fundamental philosophical "Lebensgefühl" (23); it has been ascribed to various nations (9) and races (24) in varying proportions, usually determined by the nationality or race of the writer; it has been used as an aid in classifying and diagnosing mental illness (8, 10, 20, 29, 33, 35); it has been correlated with personality and temperament (11, 19, 25, 28, 31), as well as with scholastic aptitude, emotional maturity, height, and weight (37). Yet in spite of these manifold uses of the term, scientific measurement of the trait has lagged seriously behind.

The following methods are the only ones which appear to have been employed to date in an attempt to quantify this elusive trait; none of them can escape serious criticisms on the theoretical level, and none of them can be said to have gained wide acceptance.

(1) One common method is that of determining a person's "humor" by comparing his ranking of humorous items with that of a standard group. This is the method used, for instance, in the Roback test (34).

(2) A similar method, adopted by Almack in his test (1), and also used by other investigators (4), consists in comparing the absolute judgments of funniness of various humorous items with a standard derived from a large group.

(3) The third method, used by Moss in his test (30), requires the subject to select the funniest ending for a joke, several different endings being provided.

(4) A fourth method, not yet officially embodied in a standardized test comparable to the Roback, Almack, and Moss tests, was used first by Claparède (7) and Harrower (21) in their work on the psychology of the higher thought processes, and is similar to the third method. The subject is presented with a cartoon, and is required to produce a funny caption for it, or with an unfinished joke, which he is required to finish. The present writer has attempted to develop this method into a standard test in some unpublished work.

Method (3) has been shown to provide a good test of intelligence but to show little evidence of any specific "humor" factor (38); the other "production" method, while undoubtedly also highly influenced by intelligence, appears more hopeful, but is as yet in the experimental stage. This leaves us for practical purposes with the two "appreciation" methods described above.

On closer scrutiny, these two methods involve two different principles. We may lay stress on the *cognitive* agreement between the subject and the criterion group, determining the number of times his judgment regarding the goodness or badness of a joke agrees with that of the standard group, or we may lay stress rather on the *affective* score of the subject, i.e., we may enquire whether he finds very many or very few jokes funny. Method (1) is better adapted to deal with *cognitive* agreement, method (2) is better adapted to deal with *affective* scores. The virtues of both methods can be combined by having the subjects rank the items, and by asking them at the same time how many items they find amusing or funny (11).

Before we can use any of these methods, we must attempt to answer a number of questions which cannot fail to arise. We must know whether there are any sex differences, and of what kind they may be. We must know whether a score in a test constructed on the above principles is specific to that test, or whether it has any predictive power with regard to other, similar tests. (In other words, we must know whether tests

of "appreciation" of humor tend to correlate together in the same way that intelligence tests do.) Similarly, we want to know which scores, *cognitive* or *affective,* as determined by one test, have greater predictive value with regard to the whole universe of possible tests. Without an answer to these and similar questions, scores on the standard tests mentioned must remain meaningless numbers, of unknown significance.

As hardly any of these questions have been answered by the authors of the above-mentioned tests, an attempt was made to provide the required information. Some of the questions asked can best be answered by means of an experiment using a large number of humorous items, and a small number of subjects whose temperament, personality, etc., can be thoroughly investigated. An experiment of this type, employing three tests containing 100, 52, and 37 items, respectively, and using 16 subjects who also took part in a temperament test, has been reported elsewhere (11). Some of the questions asked can best be answered by means of an experiment using a smaller number of items, and a large number of subjects. An experiment of this type is reported in the pages that follow.

The Experiment

Material. The experimental material consisted of five sets of twelve items each. The sets were named Jokes, Pictures, Limericks, Verses, and Comparisons: their nature can be deduced from these names. The items comprising four of these sets are included in the present article, appendix A, and a description of the fifth set (Pictures) is also given. The items in each set were labelled A through M, the letter J being omitted, as some subjects mix it up with the letter I.

Subjects. These five tests were given to 100 subjects, 50 male and 50 female. The average age of the male group was 31.1, the average age of the female group 28.3; the average age of the whole group was 29.7. The subjects formed a rather heterogeneous sample of the population, including at one end eleven graduate and post-graduate university students, and at the other end six Civil Defence Workers from one of London's

worst slum districts. The great majority of subjects were lower middle or working class. All were British by birth.

Procedure. The subjects were asked to rank the items in each of the five tests in order of "funniness," putting the most amusing one on top and the least amusing one at the bottom. They were instructed to judge entirely by their own subjective feeling in the matter, without trying to think what the majority opinion would be in each case. They were told that there were no right or wrong answers. In addition to ranking the items in each set, they were asked to state how many of the items they found amusing, and how many items they were familiar with.

TABLE 1

The Average and Standard Deviation of the Ranks of Each Item for Each Test

Item	Jokes		Pictures		Limericks		Verses		Comparisons	
	M	σ	M	σ	M	σ	M	σ	M	σ
A	6.5	3.80	6.2	3.29	5.2	3.33	4.6	3.36	5.0	3.38
B	7.8	3.19	4.7	2.93	8.3	3.23	6.7	3.29	7.9	3.06
C	6.4	3.43	5.3	3.85	6.2	3.34	4.9	3.26	6.0	3.29
D	5.7	3.64	7.3	3.22	6.1	2.96	8.4	3.24	8.5	3.39
E	4.2	3.12	6.0	2.96	6.8	3.29	7.4	2.70	8.7	2.86
F	6.9	3.12	5.6	3.32	5.3	3.42	7.4	3.33	6.8	3.29
G	7.8	3.20	7.3	3.29	4.4	3.18	6.8	3.25	5.6	3.23
H	7.2	3.19	7.5	3.28	6.4	3.03	7.5	3.27	5.2	2.72
I	5.9	2.93	8.9	2.83	8.1	2.77	6.2	3.32	5.3	3.29
K	6.3	3.28	5.4	3.06	8.1	3.30	6.3	3.26	8.0	2.96
L	7.2	3.31	6.9	3.39	8.2	3.18	5.4	3.66	5.0	3.32
M	6.2	3.53	6.9	3.50	4.9	3.03	6.5	3.51	6.2	3.37

Results. In Table 1 are given the average positions and the standard deviations of each of the twelve items in each of the five tests. It will be seen that the best-liked item in any test has a position of 4.2, while the least-liked item has a position of 8.9; this difference in position of 4.7 points contrasts with a maximum possible difference of 11 points. The standard deviations show considerable variation; the least variable item has a S.D. of 2.70, and the most variable item a S.D. of 3.85. The average S.D.'s of the five tests are: Jokes = 3.31; Pictures = 3.24; Limericks = 3.17; Verses = 3.29; Comparisons = 3.18. It may be concluded that there is approximately as much divergence of opinion about the "funniness" of the items in one of the five tests as there is about that of the items in any of the others.

TABLE 2

Average Numbers of Items Liked and Known

Test	Average number liked				Average number known			
	Men	Women	Total		Men	Women	Total	
	M	M	M	σ	M	M	M	σ
Jokes	3.98	4.04	4.01	2.89	1.46	.76	1.11	1.69
Pictures	5.67	5.20	5.43	2.88	3.78	1.98	2.88	4.12
Limericks	3.90	4.00	3.95	2.72	1.70	1.51	1.61	2.90
Verses	3.36	3.86	3.61	2.94	.42	.22	.32	.94
Comparisons	4.06	4.36	4.21	2.65	.82	.46	.64	1.84
Average	4.19	4.29	4.24		1.64	.99	1.31	

In Table 2 are given the average numbers of items liked ("found amusing") and known in each of the five tests. On the whole, about 35% of the items were found amusing, and about 10% were known. While the number of items liked does not differ much from test to test (particularly considering the large S.D.'s), there are great differences among the tests regarding the number of items with which the subjects were familiar. The "Pictures" were best known, the Limericks and the Jokes were much less well known, and the Comparisons and the Verses were hardly known at all. While there was hardly any difference in the number of items liked between the men and the women (4.19 and 4.29, respectively), there was a great difference in the number of items known between men and women. The men, on the average, claimed to know almost twice as many items as the women (1.64 and .99 are the respective figures).

Items which are known are liked slightly better than items not known. On the average, the known jokes rank 5.9, the known pictures 6.2, the known limericks 5.8, the known verses 6.0, and the known comparisons 5.6, as compared with a chance average of 6.5. Whether this is due to the fact that people tend to prefer familiar material, or whether better material is more likely to get known, it is impossible to say. It might even be possible that better-liked material is better remembered.

When the number of items liked in each of the tests is plotted, the resulting distribution is normal with the exception of the values for 0 (i.e., no items liked at all), which is too

large in every single test. Summing the distributions for the five tests, we get Figure 1, which clearly illustrates this phenomenon.

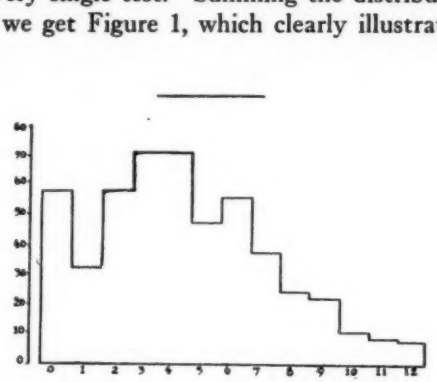


FIG. 1. Number of items liked: totals for 5 tests and 100 subjects.

TABLE 3

Average Intercorrelation Between Rankings in Each Test

Test	Average intercorrelations		
	Men	Women	Total
Jokes	.068	.067	.067
Pictures	.063	.148	.098
Limericks	.121	.174	.148
Verses	.106	.103	.090
Comparisons	.121	.165	.145
Mean	.096	.131	.110

The average intercorrelations between the rankings in each of the tests are shown in Table 3. These intercorrelations, calculated from the values in Table 1 by means of Kelley's formula (27), are surprisingly small. On the average the men's rankings correlate only to the extent of .096, and the women's to the extent of .131. The average of all the 24,750 correlations is .110. In the writer's previous investigation an average correlation of .16 was found for the rankings of 187 items, grouped in three tests, by 16 subjects. The reduction in the size of the average intercorrelation from the previous work may be accounted for by the greater heterogeneity of the population;[1]

[1] When *tests* are correlated, homogeneous populations will give smaller average intercorrelations than heterogeneous populations; when *persons* are correlated, this rule is reversed. Similarly, the more homogeneous the experimental material, the

another possible explanation, greater homogeneity of the experimental material, seems unlikely as a causal factor because the same principles of selection were employed in both cases.

As the writer has shown elsewhere (12, 13), it is possible to find from the average intercorrelations of a number of rankings the correlation of the sum of these rankings with the "true" order of merit of the items ranked, i.e., with that order which would have resulted if an infinite number of subjects had been used. These values are given in Table 4; it will be seen that

TABLE 4

Correlations of Average Rankings with "True" Rankings

Test	Correlations of average with "true"		
	Men	Women	Total
Jokes	.89	.89	.94
Pictures	.87	.94	.96
Limericks	.93	.95	.97
Verses	.93	.92	.95
Comparisons	.93	.95	.97
Mean	.91	.93	.96

the average rankings of the items in all the five tests used are highly valid; on the average, our empirical rankings correlate to the extent of .96 with the "true" rankings.

If there are no differences in the factors which make men and women prefer one item to another in the tests used, then the correlations between the average rankings of the men and the average rankings of the women should equal the product of the respective correlations with the "true" order of the two groups. These theoretical values are set out below, in Table 5,

TABLE 5

Theoretical and Actual Correlations Between Rankings of 50 Men and 50 Women

Test	Theoretical	Actual
Jokes	.79	.76
Pictures	.82	.78
Limericks	.88	.91
Verses	.86	.61
Comparisons	.88	.74
Mean	.83	.76

larger will the correlations be when *tests* are correlated, while the correlations will be smaller when *persons* are being correlated.

together with the actual correlations. It will be seen that the average theoretical correlation between men's and women's rankings is .83, as compared with an actual correlation of .76. These values agree too closely to allow us to assume any considerable difference between the appreciation of humor of men and women, respectively. This conclusion is in agreement with earlier work (11, 22).

Scores for tests of the kind employed here can be derived by correlating each person's rankings in each of the tests with the average ranking of the items in those tests by the whole group (14). We have already seen that the average rankings are highly valid, correlating to the extent of .96 with the "true" order. The distribution of the resulting 500 scores, grouped about their common average in terms of sigma units, is shown in Figure 2. It will be seen that the scores are distributed roughly in the shape of the normal distribution curve.

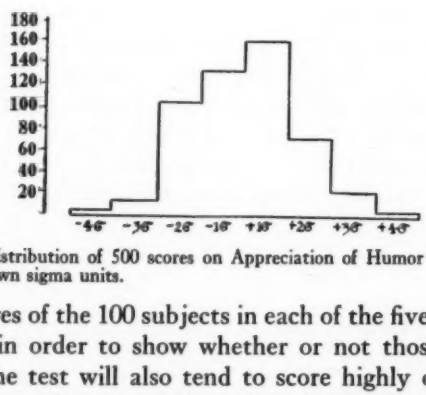


FIG. 2. Distribution of 500 scores on Appreciation of Humor tests, plotted in terms of their own sigma units.

The scores of the 100 subjects in each of the five tests can be correlated, in order to show whether or not those who score highly in one test will also tend to score highly on the other tests. In a previous paper the writer showed that the correlations between three tests of appreciation of humor were not statistically significant; similarly Stump did not find any significant correlation between the two parts of the Almack humor test (37). The correlations among the five tests employed in this study are given in Table 6; it will be seen that

TABLE 6

Correlations Among the Five Tests

Test	Jokes	Pictures	Limericks	Verses	Comparisons
Jokes	(–)	–.03	.15	–.09	.05
Pictures		(–)	–.07	–.03	.13
Limericks			(–)	.10	.18
Verses				(–)	–.02
Comparisons					(–)
Average	.02	.00	.09	–.01	.09

there is no evidence of any tendency for those who score highly in one test to score highly in any other test. The average intercorrelation is only .0426 (the original correlations were calculated to four decimals, but are given to only two decimals in the table). This result is very different from the results of similar experiments carried out in the field of aesthetics, where the writer showed that eighteen tests of aesthetic appreciation correlated together significantly, giving rise to a general factor (called "T") which accounted for 21% of the variance (14).

In a previous experiment, it had been found that people who report that they find a large number of items amusing in one test also tend to report that they find a large number of items amusing in the other tests (11). The correlations among the five tests for number of items found amusing are given in Table 7; they average .57, as compared with an average inter-

TABLE 7

Correlations of Numbers of Items Found Amusing for the Five Tests

Test	Jokes	Pictures	Limericks	Verses	Comparisons
Jokes	(–)	.57	.67	.62	.63
Pictures		(–)	.46	.42	.47
Limericks			(–)	.65	.57
Verses				(–)	.64
Comparisons					(–)
Average	.62	.48	.60	.58	.58

correlation of .55 in the previous experiment. In both sets of experiments the "Jokes" test is the most diagnostic of this tendency, although in the one case it consisted of 100 items, in the other of only 12.

A slight positive (insignificant) correlation had been found in the previous experiment between liking a large number of items, and having a high score (11, p. 299). In the present experiment the following correlations were found for the five tests: Jokes = + .079; Pictures = + .009; Limericks = + .128; Verses = + .065; Comparisons = − .036. The average of these correlations is only .049, which is clearly not significant.

Several interesting group factors had been found in the previous experiment (11, pp. 299–300), and an attempt was made to determine the presence and relative importance of group factors in the present series of rankings. The method used was slightly unusual, and is based on the theorem that with the exception of the first, or general, factor, *the factors extracted from a matrix of correlations between persons are identical with the factors extracted from a matrix of correlations between tests (or test items).*[2]

Now, the scores derived from correlating each person's ranking with the average ranking are in effect approximations to centroid factor saturations derived from the matrix of intercorrelations between the subjects. In order to determine the existence and nature of group factors we can either (a) calculate the 24,750 intercorrelations of the 100 subjects for the 5 tests, factorize the resulting matrices, extracting first of all the general, positive factors to which our scores approximate, and then re-analyze the residual matrices for group factors, or (b) we can rest content with our scores as approximations to the general factor saturations and go on to correlate the *items* in each of the tests. In accordance with the law quoted in the preceding paragraph, this latter procedure should eliminate the general factor and produce exactly the same group factors which we would have got by means of the first procedure.

As the labor involved in the first method outlined above is prohibitive, the second method was used in this research. This method, apart from being much more economical, has the added advantage that the items defining the group factors are obtained directly, and do not have to be inferred from the

[2] Full discussion, formulation, and proof of this statement are given in (5).

judgments of the subjects obtaining particularly high positive or negative saturations, as is the case in method (a). The advantage of this procedure will be realized by all who have had to decide just how many subjects should be used in order to delineate group factors, and which subjects' factor saturations should be deemed too low for them to be included in this determination.

First of all, the raw rankings were inspected carefully to discover those tests in which there might be a reasonable hope of finding significant and meaningful group factors. The "Jokes," "Pictures," and "Limericks" tests were considered the most likely to show evidence of group factors, and accordingly were analyzed by means of the method outlined above.

The resulting three twelve-square matrices contained positive and negative correlations in roughly equal proportions. When $N = 100$, a correlation of .20 is significant according to Fisher's method ($p = .05$), while a correlation of .25 is very significant ($p = .01$). The three matrices contained, respectively, 10, 12, and 12 significant correlations, and 4, 8, and 9 very significant correlations. (By chance, we might have expected 3 significant and ½ very significant correlations in each matrix.) We may conclude that all three matrices show evidence of group factors.

Factor analysis of these three matrices results in three group factors which account, respectively, for 8%, 10%, and 11% of the variance. Psychologically, these factors confirm analyses made in an earlier experiment, in which a much smaller number of subjects, a much larger number of items, and a different method of factor analysis had been used. This agreement provides welcome confirmation of the validity both of the psychological analysis and of the statistical procedure here advocated.

As regards the analysis of the "Jokes," the items seem to fall into two groups which may roughly be identified as "clever" and "funny." The clever jokes are less obvious, usually circumlocutory, and generally require the mind to jump from what is said to what is implied. The funny jokes are very obvious, deal frequently with topics pertaining to the body, and demand

no subtlety of any kind. Examples of clever jokes, as determined by the analysis, are F, D, I, M, L; examples of funny jokes are A, B, and C. While these jokes substantiate the above analysis, it should be mentioned that the analysis is partly determined by the much larger number of jokes which was found to fall into much the same pattern in the writer's earlier work (11).

The "Pictures" test also confirms a group factor found in the earlier work, viz., a factor dividing situational from personal humor, i.e., humor based on the characters of the persons depicted vs. humor based on the situation, irrespective of character. Examples of the former type of humor are C, F, and M, depending on the character qualities of tippling old women, A.T.S. officers, and vacuum cleaner salesmen. Examples of the latter type are A, B, and G, in which the characters of the persons are of little importance. Unfortunately it was impossible to reproduce the cartoons so as to allow the reader to form his own judgment of the adequacy of this distinction; the descriptions given in Appendix A are hardly adequate.

The "Limericks" fall into two groups which seem to correspond with the categories found suitable for the jokes, viz., clever vs. funny. The funny limericks are those which are just absurd and nonsensical, such as C, L, E, D, B, while the clever limericks contain some twist which appeals to the intelligence. Examples of the latter type are H, F, G, A, and M. These differences will be elaborated more fully in the discussion of "sense of humor" in the following section.

Discussion

Discussion may be clarified by seeking agreement on the nature of the factors which determine our appreciation of humor. The writer has applied Burt's four-factor theory (5) to a classification of the factors determining aesthetic appreciation (15), and it would appear that the results sanction a similar application in the field of humor.

Following in general the argument of this previous paper (15), we find first of all a *general* factor, i.e., an indication that there is some kind of agreement among *all* our subjects on the

"funniness" or otherwise of the jokes or other humorous items. When this general factor is eliminated by a suitable experimental or statistical procedure, *group* factors emerge, i.e., factors which characterize the preference judgments of groups of subjects. Such factors are, for instance, the "funny" vs. "clever" factor, which bears a resemblance to the "simple" vs. "complex" group factor which was found in the author's work in aesthetics. This factor appears to be correlated with temperament, the introverted subjects tending to prefer the "clever" and "complex" items, the extraverted subjects preferring the "funny" and "simple" items (16, 11).

While these two factors are "communal" in the sense that all or a certain number of observers agree on the ranking of the items, two further factors are "unique," i.e., they are peculiar to each of the judges separately. First, *individual judgments* may be based on personal associations and experiences. A certain amount of experimental work has been done under this heading, both in the field of aesthetics (17) and of humor (3, 26). And also we have *error factors,* i.e., factors which do not remain stable even within the same person, but which vary from day to day. Their relative contribution is reflected in the retest reliabilities.

For such samples of the population (both of subjects and of items) as were used in this experiment and in the writer's previous work, the importance of these four types of factors appears to be roughly in the ratio of 1: 1: 6: 2. In other words, while individual factors and error factors accounted for some 80% of the variance, general and group factors accounted for only about 20%. In the author's work on preference judgments in aesthetics, "communal" factors were considerably more important (15).

It is perhaps because the general factor is so much less important in the present series of tests than it was in the series of 18 tests of aesthetic appreciation used by the author in attempting to establish the general factor of aesthetic appreciation that we find no correlation between the scores in the five individual tests used. While this is a possible reason, it does not appear a very likely one; a reduction in the importance of the general

factor might result in a greater influence of chance elements and hence in a reduction in the correlation between the tests; yet it is difficult to see how such a reduction could lead to the total abolition of any correlation between the tests.

The true position would appear to be that while the factors which determine our preferences in one test of aesthetic appreciation are closely related to, or partly identical with, those which determine our preferences in other tests of the same kind, *no such similarity or identity can be found among the factors determining our appreciation of different types of humorous items.* This conclusion, perplexing though it may be, and unclear as its implications may appear, seems to be forced on us by the results described in the previous section. If this conclusion be accepted, it follows as an important consequence that scores on the ordinary type of test of "appreciation of humor" are valueless when regarded as giving an estimate of the subject's general "sense of humor." They measure nothing but the subject's reaction to the test itself.

The position is quite different when we turn to the other method of scoring, i.e., in terms of number of items found amusing. Here we find that it is possible to predicate a general factor embracing all the tests. This general factor, indeed, appears to be even stronger than that found in analyzing the correlations between the eighteen tests of aesthetic appreciation (14). This would suggest that the essence of "sense of humor" is affective rather than cognitive; or rather, that there is great individual constancy in the manner in which subjects react affectively to humorous items of any kind, but that their cognitive reactions are determined largely by factors of individual experience, and do not allow prediction from one test to another. The relevance of this finding to test construction is too obvious to be stressed. It should also be noted that our conclusion agrees well with clinical findings which have not always been taken into consideration by psychological test-constructors (6, 18, 32, 36, 39).

The results as reported above seem to strengthen the theoretical analysis of "sense of humor" given by the writer in a previous paper (11). There it was shown that three factors

demand to be taken into account in explaining the amusement derived from a joke (taking joke as the generic name for all types of humorous items). These three factors are cognitive, conative, and affective, and the three types of reaction to which they give rise were called the *Comic, Wit,* and *Humor,* respectively. An analysis of the cognitive factors showed that there are five points which enter into the joke; we may summarize them by saying that we consider something as "funny" or "amusing" when there is *a sudden, insightful integration of contradictory or incongruous ideas, attitudes, or sentiments which are experienced objectively* (11, p. 307).[3] Affective factors were found to include a state of joyfulness, while conative factors included a feeling of "superior adaptation," as Ludovici calls it.

Correlations between group factors and the results of a temperamental analysis seemed to show that preferences for "clever" jokes indicated a prevalence of cognitive factors, and an introverted temperament, while preference for "funny" jokes indicated a prevalence of conative and affective factors, and an extraverted temperament (11).[4] The results of the present study, as far as they go, are in full agreement with this analysis.

Summary and Conclusions

Five tests of "appreciation of humor," each consisting of 12 items to be ranked in order of "funniness," were given to 50 male and 50 female subjects, representing a sample of the population fairly representative as regards social class, age, etc. In addition, the number of items known and the number of items found funny were noted for each test and for each subject. Scores were established for each person in each test by correlating the subjects' rankings with the average ranking. The items in each of three tests were intercorrelated, and the result-

[3] According to the "Law of Reinstatement by Partial Content," any one member of a causal series may in time suffice to produce an effect which was at first produced only by the action of all the members of that series. This qualification of the above principle should be kept in mind when analyzing actual instances of laughter and amusement.

[4] These results confirmed Kambouropoulou's finding that "the more extraverted subjects have a greater proportion of the superiority class among the items they find most amusing. Extraversion and preference for the superiority class of humorous items go together" (26, p. 55).

ing matrices factorized, in order to discover group factors. The five tests were intercorrelated, both for scores and for "number of items liked," in order to find whether there were any factors common to all tests. The following conclusions were arrived at:

(1) There was as much divergence of opinion on the "funniness" of the items in one of the tests as there was on that of the items in any of the other tests.

(2) On the whole, about 35% of the items were found amusing, and 10% were known.

(3) Men and women did not differ with regard to the number of items liked, but the men claimed to know on the average twice as many items as the women.

(4) There was a slight tendency for items which are known to be liked better.

(5) The rankings of the 100 subjects in each of the tests intercorrelated to the extent of .110 on the average.

(6) The rankings of the women intercorrelated on the average higher than those of the men, the respective correlations being .131 and .096.

(7) The validity of the average rankings was very high; it equalled .96 on the average.

(8) There did not appear to be any great difference between the preferences of the men and the women, as shown by their average rankings.

(9) There were no significant correlations between the scores of the 100 subjects in the five tests used.

(10) There was an average intercorrelation of .57 for "number of items found amusing" among the five tests.

(11) There was no significant correlation between liking a large number of items and having a high score.

(12) Three group factors were found in an analysis of the intercorrelations of the items in three tests:

a) "Clever" jokes vs. "funny" jokes. ("Jokes" test.)

b) "Situational" vs. "character" humor ("Pictures" test.)

c) "Clever" limericks vs. "funny" limericks. ("Limericks" test.)

(13) The results of this experiment, as far as they go, confirmed the theoretical analysis of "sense of humor" given elsewhere by the writer.

APPENDIX A

Test One: Jokes

(A) Customer: "Why does that dog hang around looking at me?" Barber: "Well, sometimes I clip off a small piece of a customer's ear."

(B) "Why do you want Limburger cheese packed in your lunch?" asked the grocer. "Because, papa," answered the truthful little son, "I want teacher to send me home."

(C) Little Willie's mother had just seen him put his thumb to his nose and wiggle his fingers at his little playmates. "Willie," she cried "What do you mean by doing such a thing?" "Don't worry, mother," said Willie. "They know what I mean."

(D) There was a heavy storm at sea, and a nervous woman passenger went to the Captain. "Captain," she asked, "are we in great danger?" "Madam," he replied, "we are in the hands of God." "Oh!" she exclaimed, "Is it as bad as that?"

(E) A woman visitor to the London Zoo asked the keeper whether the hippopotamus was a male or a female. "Madame," replied the keeper sternly, "that is a question which should be of interest only to another hippopotamus."

(F) Associate Editor of Humorous Paper: "Let's not print any more Scotch, Jewish, Ford, or Hitler jokes." Editor: "All right. I'm tired of putting out this lousy magazine, anyway."

(G) "You are the sunshine of my life. Your smile drives every cloud away. With you at my side, I would defy the storms of life." "What is this? A proposal or a weather report?"

(H) Hubby: "One night while you were away I heard a burglar. You should have seen me coming down the stairs three at a time." Wifie: "Where was he? On the roof?"

(I) Prison Visitor: "How long are you here for?" Convict: "Thirty years." Visitor: "Ah well, here's another day nearly gone."

(K) Old lady: "I wouldn't cry like that, my little man."
Boy: "Cry as you please, this is my way."

(L) A man living in a village outside Paris during the Revolution met a friend fresh from the city and asked what was happening. "It's awful," was the reply; "they're cutting off heads by the thousand." "Good Heavens! Surely not heads," he cried. "Why, I'm a hatter."

(M) The speaker, who had arrived in a crabby frame of mind, looked around and beckoned the chairman. "I would like to have a glass of water on my table, if you please," he said. "To drink?" was the chairman's idiotic question. "Oh, no," was the sarcastic retort; "when I've been speaking half an hour I do a high dive."

Test Two: Pictures

(A) A castle, with wings, flying through the air. One yokel pointing it out to the other, saying: "Look, a flying fortress. That'll show 'em."

(B) Soldier with camouflage branches in his helmet, reporting to his officer. A bird has made its nest in the branches.

(C) Old woman, looking at remains of bombed house. Sees whisky bottle inside battered old gramophone, says: "I always wondered where she kept it."

(D) Oxford don, writing on the college wall with a piece of chalk: "Septimus Perks M.A. is dementia praecox."

(E) Policeman, pointing out a sign "Diversion" to a cyclist who was riding past it: "What the devil d'you think we put that there for? A diversion or something?"

(F) Witch has turned vacuum cleaner salesman into a frog, saying: "And if the cleaner does all you say, I will remove the spell."

(G) Maid, holds the head of a dog in her hand while the headless body stands behind her: "He's been barking his head off."

(H) Goering offers Hitler some alcohol at a party. "Well, just the teeniest bit, as it's Christmas—but I mustn't get into wrong ways."

(I) Ghost holds his head in amazement as owner of castle walks along the corridor with his head under his arm.

(K) Hitch-hiker signalling bomb-disposal car with live bomb in it.

(L) Air-raid wardens in front of shelter, talking about a third warden: "Weird sort of cove: each time a bomb drops he says 'that's 'andy.' "

(M) Contrite A.T.S. officer being addressed by superior officer in front of other A.T.S.: "For conduct unbecoming to an officer and a gentlewoman, Captain Winkworth will be 'kept in' for a week."

Test Three: Limericks

(A) There was an old party of Lyme,
Who married three wives at one time,
When asked, "Why the third?"
He replied, "One's absurd,
And bigamy, sir, is a crime."

(B) There was a young man of Laconia,
Whose mother-in-law had pneumonia;
He hoped for the worst,
And after March first
They buried her 'neath a begonia.

(C) There was a young girl of Asturias,
Whose temper was frantic and furious.
She used to throw eggs
At her grandmother's legs
A habit unpleasant, but curious.

(D) There was an old man of Tralee
Who was bothered to death by a flea;
So he put out the light,
Saying, "Now he can't bite,
For he'll never be able to see."

(E) There was a young lady of Ealing,
Who had a peculiar feeling
That she was a fly,
And wanted to try
To walk upside down on the ceiling.

(F) There was a young woman named Bright,
Whose speed was much faster than light,
She set out one day,
In a relative way,
And returned on the previous night.

(G) There was a young lady named Starkey,
Who foolishly married a darkey.
And then for her sins,
She had three pairs of twins,
One white, one black, and one khaki.

(H) For beauty I am not a star,
There are others more handsome by far;
But my face I don't mind it,
For I am behind it,
It's the people in front that I jar.

(I) There's a very mean man of Belsize,
Who thinks he is clever and wise.
And what do you think?
He saves gallons of ink
By simply not dotting his "i's."

(K) A prideful young lady of Boston
A two-horned dilemma was tossed on;
As to which was the best,
To be rich in the West
Or poor and peculiar in Boston.

(L) There was an old soldier of Bister
Went walking one day with his sister;
When a cow at one poke
Tossed her into an oak
Before the old Gentleman missed her.

(M) An epicure, dining at Crewe,
Found quite a large mouse in his stew
Said the waiter, "Don't shout,
And wave it about,
Or the rest will be wanting one, too."

Test Four: Verses

(A) Here lies the body of poor old Jones,
Who all his life collected bones.
Until came death, that mighty spectre,
That all-amassing bone collector,
And boned old Jones so neat and tidy,
That here he lies, all *bona fide*.

(B) Lady, lady, should you meet
One whose ways are all discreet,
One who murmurs that his wife
Is the lodestar of his life,
One who keeps assuring you
That he never was untrue,
Never loved another one—
Lady, lady, better run.

(C) How courteous is the Japanese,
He always says, "Excuse me please."
He climbs into his neighbour's garden
And smiles and says, "I beg your pardon."
He bows and grins a friendly grin,
And calls his hungry family in;
He grins and bows a friendly bow:
"So sorry, this *my* garden now."

(D) Or ever a lick of Art was done,
Or ever a one to care,
I was a Purple Polygon,
And you were a Sky-Blue Square.
I yearn for you, but I have no chance,
You lie in a different plane,
I break my heart for a single glance,
But I break said heart in vain.

(E) The Zebra, born both black and white,
Is just the jungle clown,
The lionesses hunt him up,
The lions hunt him down.
His life, in consequence is brief
And seems inclined to end in grief.
And so, you see, between the two,
He's more contented in the zoo.

(F) This cheerful story tells the fate
Deserved by William Smith—the late—
When he had killed two lions, he
Was killed in turn, by number three.
The lion's notable behavior
Was printed in the roto-gravure.

(G) The crocus grows in any spot,
And multiplies an awful lot.

It doesn't pout, and fail to bloom,
Because of soil, or lack of room.
No books are written on the crocus;
It grows without such hocus-pocus.

(H) Marco Polo travelled far,
Went through perils worse than war,
Saw great marvels, distant lands,
Unicorns and sarabands.
Why, and wherefore? So that when
He had toiled back home again,
He could chant an endless solo
On the deeds of Marco Polo.

(I) Down in the silent hallway
Scampers the dog about,
And whines, and barks, and scratches,
In order to get out.
Once in the glittering starlight,
He straightway doth begin
To set up a doleful howling
In order to get in.

(K) The Horse is not supposed to know
How to reap or how to sow;
How to boast or how to bind
Dictionaries in the Mind;
How to build a rabbit-hutch—
But it doesn't matter much,
For he understands, of course,
Exactly how to be a horse.

(L) Gorillas, with intent to please
The simple-minded, bend their knees
And walk through equatorial lands
Supported on their horny hands.
—A practice, it is fair to add,
That many now admit is bad
And only have recourse to in
Emergency, or after gin.

(M) I recollect a nurse call'd Ann,
Who carried me about the grass,
And one fine day a fine young man
Came up, and kiss'd the pretty lass.
She did not make the least objection;
Thinks I, "Aha."
"When I can talk I'll tell Mamma."
—And that's my earliest recollection.

Test Five: Comparisons

(A) He came into the room like a squirt from a syphon.
(B) He took in the room with a glance like a lasso.

(C) He adjusted his face as though it were a necktie.
(D) A croquette is nothing but a hash that has come to a head.
(E) My heart was going like a woodpecker at my side.
(F) He had no more initiative than a shadow.
(G) She approached with the slow dignity of a ferry coming into dock.
(H) He was as inconsiderate as an alarm clock.
(I) She was as maternal as an incubator.
(K) He was as irresponsible as a streak of lightning.
(L) She turned him down like a bedspread.
(M) The hours crawled by like paralytic centipedes.

REFERENCES

1. Almack, J. C. *Sense of Humor Test* (Form I). Cincinnati: C. A. Gregory, 1928.
2. Allport, G. W. *Personality*. London: Constable, 1938.
3. Barry, H. "The Role of Subject-Matter in Individual Differences in Humor." *Journal of Genetic Psychology*, XXXV (1928), 112–128.
4. Bird, G. "An Objective Humor Test for Children." *Psychological Bulletin*, XXII (1925), 137–138.
5. Burt, C. *The Factors of the Mind*. London: University of London Press, 1940.
6. Casella, P. *Contribution a l'Étude du Rire et du Pleurer Spasmodiques*. Lyons, 1902.
7. Claparede, E. "La Genèse de l'Hypothése." *Archives de Psychologie*, XXIV (1934), 1–154.
8. Coriat, I. H. "Humor and Hypomania." *Psychiatric Quarterly*, XIII (1939), 681–688.
9. Egner, F. "Humor und Witz unter strukturpsychologischem Gesichtspunkt." *Arch. f. d. ges. Psychol.*, LXXXIV (1932), 330–372.
10. Elste, R. "Über die Anwendbarkeit eines Bildserientests zur Beurteilung psychischen Krankheitsgeschehens." *Ztsch. ges. Neurol. Psychiat.*, CLXX (1940), No. 1.
11. Eysenck, H. J. "The Appreciation of Humour: An Experimental and Theoretical Study." *British Journal of Psychology*, XXXII (1942), 295–309.
12. ————. "The Validity of Judgments as a Function of the Number of Judges." *Journal of Experimental Psychology*, XXV (1939), 650–654.
13. ————. "The Validity and Reliability of Group Judgments." *Journal of Experimental Psychology*, XXIX (1941), 427–434.
14. ————. "The General Factor in Aesthetic Judgments." *British Journal of Psychology*, XXXI (1940), 94–102.
15. ————. "The Experimental Study of the 'Good Gestalt'—A New Approach." *Psychological Review*, XLIX (1942), 344–364.

16. ————. "Abnormal Preference Judgments as 'Complex' Indicators." *American Journal of Orthopsychiatry*, XII (1942), 338–345.
17. ————. "Some Factors in the Appreciation of Poetry, and Their Relation to Temperamental Qualities." *Character and Personality*, XI (1940), 160–167.
18. Fyodoroff, M. I. (On irresistible laughter in some forms of nervous affections.) Vestnik klin. i subebuoi psychiat. i nevropatol., VIII (1890–91), 78–97. Abstracted in: *Ann. de psychiat. et d'hypnol.*, II (1892), 87.
19. Gregg, A., Miller, M. & Lindon, E. "Laughter Situations as an Indication of Responsiveness in Young Children." In: *Some New Techniques for Studying Social Behavior*. New York: D. S. Thomas, 1929.
20. Haggard, E. A. "A Projective Technique Using Comic Strip Characters." *Character & Personality*, X (1942), 289–295.
21. Harrower, M. "Organization of Higher Mental Processes." *Psychol. Forsch.*, XVII (1932), 56–120.
22. Heim, A. "An Experiment on Humour." *British Journal of Psychology*, XXVII (1936), 148–161.
23. Höffding, H. *Humor als Lebensgefühl*. Leipzig: Teubner, 1918.
24. Kadner, S. *Rasse und Humor*. Munich: Lehmann, 1939.
25. Kambouropoulou, P. "Individual Differences in the Sense of Humor." American Journal of Psychology, XXXVII (1926), 288–297.
26. ————. "Individual Differences in the Sense of Humor and Their Relation to Temperamental Differences." *Archives of Psychology*, 1930, No. 121.
27. Kelley, T. L. *Statistical Methods*. New York: Macmillan, 1924.
28. Landis, C. and Ross, J. "Humor and its Relation to Other Personality Traits." Journal of Social Psychology, IV (1933), 156–175.
29. Mayer-Gross, W. "Über Spiel, Scherz, Ironie und Humor in der Schizophrenie." *Ztsch. f. d. ges. Neurol. und Psychiat.*, LXIX (1921), 332–353.
30. Moss, F. A., et al. *George Washington Test of Social Intelligence*. 1926.
31. Omwake, L. "A Study of Sense of Humour: Its Relation to Sex, Age, and Personal Characteristics." Journal of Applied Psychology, XXI (1937), 688–704.
32. Oppenheim, H. "Über Lachschlag" (Lach-schwindel; Lach-Ohnmacht.). *Monatsch. f. Psychiat. und Neurol.*, XI (1902), 241–247.
33. Rebatel, A. *Le Rire dans la Démence Précoce*. Lyons: 1908.
34. Roback, A. A. *Sense of Humor Test*. Cambridge, Mass., 1939.
35. Senise, T. *Il Riso in Fisio-Patologia*. Naples: Bibl. d. Stud. 1941.

36. Sikorski, I. A. (Uncontrollable laughter and crying.) *Protok. Kiev Psikhiat. Obst.*, II (1899–1900), 13–20.
37. Stump, N. "Sense of Humor and its Relationship to Personality, Scholastic Aptitude, Emotional Maturity, Height and Weight." Journal of General Psychology, XX (1939), 25–32.
38. Thorndike, R. L. and Stein, S. "An Evaluation of the Attempts to Measure Social Intelligence." Psychological Bulletin, XXXIV (1937), 275–285.
39. Wiesner, A. (The pathology of laughter.) *Casap. lék. česk.*, 42, 1903, 141, 161, 201, 226, 251.

THE DEVELOPMENT AND EVALUATION OF A MEASURE OF COUNSELING INTERVIEW PROCEDURES

Part II

The Evaluation

E. H. PORTER, JR.

Merit System Council for the Oregon State Public Welfare Commission[1]

In the first section[2] of this report there was described the development of a check list of procedures suitable for use in identifying procedures used in counseling interviews. The development was based upon a theoretical analysis of publications on how to counsel and an empirical analysis of phonographically recorded interviews. In this section of the report there are described the steps taken in the evaluation of the check list and the outcomes of the evaluation.

Selection of Interviews

The nineteen interviews studied were selected from phonographically recorded interviews gathered by the Educational Psychology Laboratory and the Psychological Clinic at the Ohio State University during the Autumn and Winter quarters of 1940–41.[3] The majority of the interviews were between students enrolled in a remedial study course and the counselors assigned them. The interviews, usually a half hour in duration, covered personal as well as academic problems. In accordance with the policy of the laboratory, the students were

[1] On military leave of absence.

[2] Porter, E. H., Jr., The Development and Evaluation of a Measure of Counseling Interview Procedures: Part I, The Development. *Educational and Psychological Measurement,* III (1943), 105–126.

[3] For a complete discussion of the recording program, the apparatus used, difficulties encountered, and transcription methods, see Covner (1941). (Bibliography contained in first section of this report.)

all told that their interviews might be overheard. Much was done to lessen distraction by concealing the microphones in gooseneck lamps placed on the table between student and counselor. The recorder was in a distant room.

The recordings were selected on the basis of five criteria:[4] (1) sufficient clarity of the recording to assure reasonable audibility; (2) equal sampling of early, middle, and late interviews within interview series; (3) a sampling of extent of experience on the part of the counselors ranging from one month to several years; (4) a sampling of interviewers differing in counseling philosophy; and (5) approximately equal duration of the interviews. One interview extended for 50 minutes.

For one portion of the evaluation the interviews were used in their phonographic or recorded form in a manner to be described later. For another portion of the evaluation the interviews were used in the form of typescripts of the recordings.

Selection of Judges

The selection of judges, the persons who were to use the check list in identifying the procedures in the interviews, was limited to persons who upon the basis of their training and experience would be acknowledged to be professionally competent yet who ranged from rather limited to quite extensive experience. The precaution was taken to assure that differing counseling philosophies were represented among the judges just as among the counselors. It might be well to note that a recorded interview conducted by the writer was among the nineteen studied and was rated as one of the least directive.[5] As the writer emphasizes a less directive approach in many of his interviews, he has tried at every point to institute suitable controls.

Data-Gathering Procedures

Since the judges were unfamiliar with the check list at the outset, experience in its use had to be provided. First, a judge

[4] Complete information regarding the recordings selected may be obtained from the writer's dissertation bearing the same title as these reports. The dissertation is on file with the library at Ohio State University.

[5] In the first section of this report directiveness was defined as the extent to which the counselor imposes upon the interview the direction which it takes.

was given written instructions describing the general outline of the activities in which he was to engage, written illustrations of each check list category, and a sample interview in typewritten form with the counselor responses identified on the margins of the pages. Following the written instructional material each judge worked with two typescripts, coding on the left margins each identified response. The judge and writer then compared their codings and attempted to clarify the meaning of each category. After two additional typescripts the judge and writer again conferred. At the conclusion of each typescript the judge was to indicate on an eleven-point scale his judgment of the directiveness of the interviewing counselor without going back over the interview in any way.

When the judge had completed the preliminary training he used the check list on four phonographically recorded interviews. As he listened he classified each identified procedure of the counselor by entering a tally in the appropriate space before the category on a mimeographed copy of the check list.[6] When a recording had been completed by a judge he indicated his evaluation of the directiveness of the interviewing counselor by encircling the appropriate number in the scale given at the bottom of the check list. There were no particular controls on the auditory conditions except to assure reasonably adequate perception of the speech.

Both the typescripts and recordings were so randomized in presentation that each judge worked with interviews that had been conducted by counselors of differing viewpoint and with interviews from the early, middle, and late periods within interview series. Although each interview was worked on by two judges in addition to the writer, no two judges worked with more than one interview in common except in the case of the first typescript and first recording, which all judges did alike. The writer took the precaution of spacing his work with the two forms of each interview no less than a week apart. Meanwhile, of course, work on other interviews had been interposed.

[6] The check list and category codings are reproduced in this section for the reader's convenience.

Code No. ..... Judge..... Date.....

DEFINING THE INTERVIEW SITUATION

(1*a*) Defines in terms of diagnostic/remedial purposes, procedures, etc.

(1*b*) Defines in terms of client responsibility for directing the interview/reaching decisions, etc.

(1*u*) (Unclassifiable)

BRINGING OUT AND DEVELOPING THE PROBLEM SITUATION

Uses lead which:

(2*a*) Forces choosing and developing of topic upon client.

(2*b*) Indicates topic but leaves development to client.

(2*c*) Indicates topic and delimits development to confirmation, negation, or the supplying of specific items of information.

(2*u*) (Unclassifiable)

DEVELOPING THE CLIENT'S INSIGHT AND UNDERSTANDING

Responds in such a way as to indicate:

(3*a*) Recognition of subject content or implied subject content.

(3*b*) Recognition of expression of feeling or attitude in immediately preceding verbal response(s).

(3*c*) Interpretation or recognition of feeling or attitude not expressed in immediately preceding verbal response(s).

(3*d*) Identifies a problem, source of difficulty, condition needing correction, etc., through test interpretation, evaluative remarks, etc.

(3*e*) Interprets test results *but not* as indicating a problem, source of difficulty, etc.

(3*f*) Expresses approval, disapproval, shock, or other personal reaction in regard to the client (but not to identify a problem).

(3*u*) (Unclassifiable)

(4) Explains, discusses, or gives information related to the problem or treatment.

SPONSORING CLIENT ACTIVITY/FOSTERING DECISION MAKING

Proposes client activity:

(5*a*) Directly or through questioning technique.

(5*b*) In response to question of what to do.

Influences the making of a decision by:

(5*c*) Marshaling and evaluating evidence, expressing personal opinion, persuading pro or con.

(5*d*) Indicates decision is up to client.

(5*e*) Indicates acceptance or approval of decision.

Reassures (5*f*) (5*u*) (Unclassifiable)
(I) Irrelevant.
(OU) Otherwise unclassifiable. Total number of checks ———
Non-directive — 1 2 3 4 5 6 7 8 9 10 11 — Directive

Assigning Directiveness-Nondirectiveness Weights to the Categories

It will be recalled that as each judge completed his classification of procedures within an interview, he was asked to rate the interview on an 11-point scale of directiveness. An independent measure of each interview's directiveness was sought by the method of assigning to each category on the check list a weight denoting degree of directiveness. Multiplying the frequency with which a category had been used by the assigned weight, and dividing this product by the total number of identifications yields a mean directiveness value for an interview which is independent of a judge's evaluation.

These weights were assigned to the categories by a paired comparisons method. Four persons, including the writer, all of whom had participated as judges, did the pairings as a committee. Care was taken to select judges not in precise agreement in viewpoint on counseling. Agreement between three committee members determined which category in a pairing was marked as the more directive. The categories for unclassifiable items were not considered. The resulting assigned weights as reproduced below in Table 1 represent the number of times each category was judged as more directive.

TABLE 1
Directiveness Weights Assigned to Check List Categories

Code	Wt.	Code	Wt.	Code	Wt.	Code	Wt.	Code	Wt.
1a	10	2a	0	3a	1	4	13	5a	16
1b	5	2b	5	3b	2			5b	15
		2c	11	3c	7			5c	17
				3d	12			5d	2
				3e	7			5e	1
				3f	11			5f	8

The fact that the weights were assigned by persons who had acted as judges and at a time following their participation as

judges should be noted. There is the possibility that their judgments of directiveness were influenced thereby and the weightings caused to reflect unduly the particular interviews used in the study.

Reliability in Use of the Check List

Evidence of reliability in the use of the check list may be obtained (1) by comparing the number of times judges agreed in their identification of individual procedures, (2) by comparing the manner in which the identifications made by the judges were distributed throughout the check list categories and (3) by comparing the extent to which such distributions of identifications were the same for typescripts as for recorded forms of interview presentation.

Analysis shows that the judges made in all 4638 separate identifications or codings on the margins of the typescripts.[7] As there were 24 categories in the check list, it would be expected that any given counselor response by chance would be assigned a specific code by one judge once in 24 times, by two judges once in 576 times, and by three judges once in 13,824 times. Translated into per cent expectancy, these values become 4%, 0.2%, and 0.007%, respectively. Actually in 488 instances all three judges agreed in their coding, a total of 1464 or 31.6% of the codings. This agreement is roughly 4500 times greater than would be expected by chance alone.

A second index of reliability was developed by measuring the agreement between pairs of judges. Since three judges worked with each interview (except the practice interview, with which all worked) the first judge could be paired with the second, the first with the third, and the second with the third. From 18 interviews 54 pairings were available. This second index was developed as a measure which included codings that were in general agreement as well as codings in exact agreement. This was achieved by involving the number of counselor responses coded in the same general group, the 1's, 2's, 3's, 5's, and Irrelevant or Otherwise Unclassifiable. These mea-

[7] This does not include the codings for interview IV–8, the interview with which all judges worked for initial practice.

sures, called Agreement Indices (AI),[8] serve not only as indices of agreement but also serve to show the range in agreement which may be expected when but two judges are involved. They are presented in Table 2.

TABLE 2*

Agreement Between Pairs of Judges in Coding Typescripts Expressed in Per Cent as an Agreement Index

Interview number	Judge pair	AI	Interview number	Judge pair	AI
I-1	X-H	40	IV-2	X-K	62
	H-F	47		K-A	54
	X-F	58		X-A	64
I-5	X-D	59	IV-4	X-E	55
	D-A	55		E-H	37
	X-A	55		X-H	46
I-8	X-K	42	IV-6	X-C	58
	K-J	58		C-D	54
	X-J	43		X-D	60
II-1	X-L	61	V-5	X-L	55
	L-J	58		L-K	51
	X-J	64		K-K	52
II-4	X-D	64	VI-6	X-E	58
	D-B	61		E-F	55
	X-B	53		X-F	48
II-8	X-H	45	VII-1	X-I	52
	H-G	44		I-G	42
	X-G	51		X-G	52
III-1	X-E	55	VIII-1	X-A	54
	E-C	42		A-B	27
	X-C	43		X-B	31
III-4	X-F	47	IX-4	X-I	61
	F-G	47		I-J	47
	X-G	40		X-J	62
III-7	X-L	42	X-6	X-C	57
	L-I	37		C-B	51
	X-I	50		X-B	57

* Interview IV-8 is not included in this analysis since all judges did the typescript of this interview as a practice exercise.

A third indication of reliability in the use of the check list with the typescripts was obtained by measuring the agreement between the judges as to the total number of identifications

[8] The agreement index was derived according to the formula

$$AI = \frac{2(S + G_1/n_1 + G_2/n_2 + \ldots + G_8/n_8)}{T_1 + T_2}$$

in which *AI* is the agreement index, *S* the number of codings in strict agreement, *G* the number of other codings falling within the same general area, but not in strict agreement, *n* the number of categories within the general area, and *T* the total number of codings by one judge.

each made for each interview. Again the judges were grouped in pairs. A coefficient of correlation was determined for the 54 pairs of total scores. This coefficient equalled 0.95. A comparable coefficient was also derived for the recorded form of interview presentation by similarly correlating the agreement between paired judges on the total number of identifications made while listening to the records. This coefficient equalled 0.96.

A second approach to the reliability in the use of the check list was made through comparisons of the manner in which the judges' identifications were distributed throughout the check list categories. Profiles were developed by placing on the abscissae of graphs the check list category codings and on the ordinates the frequency with which each category was used to identify the counselor's responses in an interview. Edgerton, Bordin, and Molish (1941) have indicated that comparability of two profiles may be measured by correlational methods. As will be observed later the profiles obtained are often strikingly characterized by the very high frequency of two or three categories. As it was felt the correlational method adopted should be sensitive to these extremes in values, the Pearson method of correlation was used. Before the results of the comparisons between these profiles are presented, it should be pointed out that the coefficients obtained may not be interpreted as having any predictive significance. For this reason the coefficients are termed Profile Comparability Indices (PCI) to help the reader keep the distinction in mind that the indices measure comparability between profiles only.

In Table 3 are presented the Profile Comparability Indices for each of the three possible pairings of the judges on each interview both for the typescript and recorded forms. The PCI values were averaged by transmuting each value into a z value, averaging the z values and retransmuting the average z value back into a PCI value after the method of averaging coefficients of correlation described by Lindquist (1940). For the typescript form the average PCI value was 0.85 and for the recorded form, 0.84.[9]

[9] As the writer worked with each interview in typescript and recorded form it was possible to correlate the two profiles resulting from his identification for each

TABLE 3

Agreement Between Interview Profiles For Individual Pairs of Judges

Interview No.	Judge pair (typescript)	PCI	Judge pair (record)	PCI	Interview No.	Judge pair (typescript)	PCI	Judge pair (record)	PCI
	X–H	0.69	X–D	0.69		X–K	0.95	X–I	0.97
I-1	H–F	0.79	D–B	0.67	IV-2	K–A	0.91	I–G	0.64
	X–F	0.82	X–B	0.81		X–A	0.95	X–G	0.65
	X–D	0.73	X–L	0.75		X–E	0.91	X–D	0.70
I-5	D–A	0.69	L–I	0.66	IV-4	E–H	0.88	D–K	0.78
	X–A	0.84	X–I	0.69		X–H	0.86	X–K	0.52
	X–K	0.49	X–F	0.52		X–C	0.65	Avg.	
I-8	K–J	0.74	F–G	0.82	IV-6	C–D	0.68	of	0.73*
	X–J	0.53	X–G	0.68		X–D	0.95	all	
	X–L	0.94	X–F	0.86		Avg.		X–I	0.86
II-1	L–J	0.88	F–H	0.75	IV-8	of	0.65*	I–J	0.95
	X–J	0.91	X–H	0.59		all		X–J	0.88
	X–D	0.95	X–A	0.97		X–L	0.84	X–H	0.91
II-4	D–B	0.89	A–J	0.94	V-5	L–K	0.92	H–G	0.85
	X–B	0.90	X–J	0.97		X–K	0.93	X–H	0.74
	X–H	0.89	X–D	0.80		X–E	0.93	X–A	0.92
II-8	H–G	0.78	D–C	0.75	VI-6	E–F	0.95	A–B	0.72
	X–G	0.92	X–C	0.94		X–F	0.93	X–B	0.72
	X–E	0.82	X–K	0.53		X–I	0.79	X–E	0.87
III-1	E–C	0.70	K–A	0.82	VII-1	I–G	0.79	E–C	0.92
	X–C	0.53	X–A	0.92		X–G	0.61	X–C	0.83
	X–F	0.80	X–C	0.93		X–A	0.78	X–L	0.86
III-4	F–G	0.90	C–B	0.86	VIII-1	A–B	0.54	L–J	0.82
	X–G	0.84	X–B	0.94		X–B	0.52	X–J	0.93
	X–L	0.78	X–H	0.67		X–I	0.92	X–F	0.67
III-7	L–I	0.67	H–E	0.74	IX-4	I–J	0.93	F–E	0.72
	X–I	0.85	X–E	0.84		X–J	0.97	X–E	0.94
						X–C	0.95	X–L	0.91
					X-6	C–B	0.82	L–K	0.85
						X–B	0.93	X–K	0.89

*An average intercorrelation after the formula developed by Edgerton and Toops (1928).

Thus far it has been shown that (1) when absolute agreement between judges is measured by their identical codings on typescript margins, (2) when broader agreement is measured by including codings in general agreement, (3) when overall

interview. The PCI values obtained were as follows: I–1, 0.92; I–5, 0.85; I–8, 0.85; II–1, 0.93; II–4, 0.98; II–8, 0.90; III–1, 0.87; III–4, 0.98; III–7, 0.82; IV–2, 0.98; IV–4, 0.90; IV–6, 0.89; IV–8, 0.82; V–6, 0.82; VI–5, 0.67; VII–1, 0.94; VIII–1, 0.73; IX–4, 0.98; X–6, 0.93. Such comparisons suggest that the form of interview presentation may not be a major determinant of judgment during the use of the check list. Such comparisons further suggest that one person may become sufficiently consistent as to participate in measures of job performance for in-service merit ratings applied to casework positions and counselor-training programs.

agreement is measured in terms of total number of identifications made, and (4) when profiles of identifications are compared there is in each instance evidence that the check list was used with sufficient reliability to permit analytical comparisons between the identifications in one interview and those of another. Stated in different terms, the data show that the check list may be used with sufficient reliability to permit its use as a research instrument.

From the writer's observations throughout the study it is suggested that several factors may have contributed in some measure to the margin of unreliability. The data taken from the typescripts represent data taken during the training period of the judges. Had the judges been trained on other typescripts first, many disagreements due to gross misunderstandings probably would have been avoided. Further, were the categories more nearly mutually exclusive, less disagreement probably would have resulted. Some judges learned more rapidly than others. There is the additional possibility that individual styles of expression among the counselors rendered identifications more difficult for the judges. There is also the possibility that as a judge passed from an interview of less directiveness to one of greater directiveness or vice versa identifications were easier or more difficult than when he passed from one interview to another of nearly equal or equal directiveness. Still one other possibility is that the judges may have identified what they themselves would have done, although it seems unlikely in view of the fact that judges of different counseling viewpoints agreed closely in their identifications. These possibilities were all present and perhaps others as well. They form a portion of the problems to be met in further research.

Validity in the Use of the Check List

In discussing the validity of the measures of interviewing procedures obtained in this study there should be borne in mind the manner in which the check list was formulated. It will be recalled that the check list arose out of the analysis of the writings of several authors on counseling procedures as well as the empirical analysis of a number of phonographically recorded

interviews. To the extent that (1) the procedures isolated by these analyses really exist, (2) the descriptive statements forming the check list categories characterize the procedures isolated, and (3) the judges mastered the concept embodied in each category, validity in the use of the check list was assured.

Evidence that the identifications made through the medium of the check list were valid may be inferred from the directiveness-nondirectiveness scores. It seems reasonable to assume that the evaluation of directiveness made by a judge after having worked with an interview was determined by the procedures the counselor used in conducting the interview. It also seems reasonable to assume that the committee of trained judges who effected the assignment of directiveness-nondirectiveness weights to the check list categories were making their judgments upon essentially the same basis when acting as the committee as when working with the interviews.[10] If these two assumptions are correct, then it may be argued that correlation between judge-assigned directiveness scores and derived directiveness scores would indicate the extent to which the procedures within the interview were comparably identified by the judges. On the other hand, if the procedures which influenced the judges' evaluations of the interviews actually had not been identified by those judges, the subsequently derived directiveness scores should show but chance agreement with the judge-assigned directiveness scores. The coefficient of correlation between averaged judge-assigned and averaged derived directiveness scores equalled 0.94 for the typescript form and 0.93 for the recorded form.

Table 4 presents for each of the nineteen interviews: (1) the interview number, a number which identifies the counselor with a Roman numeral and the position of the interview with

[10] The reliability of the assignment of directiveness-nondirectiveness scores by judges to typescript and recorded interviews was measured by pairing the judges for each form and correlating the scores. The coefficient of reliability for assignments to typescripts equalled 0.82, and for assignments to recordings 0.86. Whatever the factors were that determined the judgments, those factors were reflected with an acceptable degree of consistency for the present study.

As the writer rated each interview in each form it is possible to determine the agreement in scores assigned under the two conditions. The coefficient of correlation equalled 0.99. This type of measure has significance also for the rating of job performance for in-service ratings on casework positions and in counselor-training programs.

TABLE 4

PCI Values, Directiveness Ratings, and Averaged Frequencies of Checks for Recorded and Typescript Forms

Interv. No.	PCI	Form	Direct. rating A.J.	Direct. rating D.S.	Codings corresponding to check list categories																										
					1a	1b	1u	2a	2b	2c	2u	3a	3b	3c	3d	3e	3f	3u	4	5a	5b	5c	5d	5e	5f	5u	OU	I
I–1	.92	R	4.3	5.7	0	3	0	1	9	9	1	7	6	4	0	0	1	0	1	2	0	0	2	1	0	0	1	0
		TS	2.0	4.2	0	4	0	0	11	6	0	7	10	6	0	0	0	0	0	1	0	0	3	0	0	0	1	1
I–5	.75	R	2.7	5.5	0	0	0	0	9	3	0	7	8	6	0	0	1	2	4	1	0	0	2	0	0	0	2	1
		TS	3.3	7.0	1	0	0	0	7	3	0	6	6	14	1	0	2	0	2	5	0	1	1	2	0	0	1	1
I–8	.93	R	1.0	5.9	0	0	0	1	7	3	0	8	8	10	0	0	1	1	5	1	0	0	1	1	0	0	3	0
		TS	2.0	5.3	0	0	0	0	6	3	0	7	11	8	0	0	1	0	4	1	0	1	0	3	1	0	2	1
II–1	.84	R	11.0	9.9	12	0	0	0	11	43	0	9	1	0	15	2	2	0	24	11	0	3	1	0	2	0	1	1
		TS	10.3	10.5	12	2	0	0	16	44	2	8	1	0	15	7	4	0	22	14	0	3	1	1	2	0	2	1
II–4	.96	R	11.0	11.3	4	0	0	9	16	45	1	3	1	0	13	0	2	0	35	18	0	6	0	0	1	0	1	1
		TS	10.7	11.9	2	0	1	0	11	56	4	3	1	0	8	1	3	0	25	15	0	6	0	1	1	1	2	1
II–8	.96	R	11.0	10.2	0	0	0	1	16	38	0	15	1	0	1	0	2	0	21	16	1	8	0	0	1	0	2	1
		TS	9.3	10.2	0	0	0	0	12	31	1	6	2	0	0	0	2	0	26	13	1	5	0	0	0	0	2	0
III–1	.95	R	10.3	8.0	2	1	1	2	17	26	0	16	1	3	2	1	6	1	32	10	0	6	0	1	2	0	4	6
		TS	9.3	9.3	2	0	0	1	22	28	1	14	1	1	2	0	2	1	23	6	1	3	1	1	3	0	4	2
III–4	.97	R	10.3	10.1	1	0	0	1	13	9	1	12	1	2	6	0	4	4	43	7	1	5	0	1	1	1	0	1
		TS	10.0	10.3	0	0	0	2	9	12	1	14	0	1	2	0	3	1	36	6	1	4	0	0	0	0	2	7
III–7	.88	R	9.7	11.8	2	0	1	1	4	4	0	3	1	1	5	0	2	0	16	6	0	11	0	0	2	3	1	3
		TS	9.3	10.7	0	0	1	2	3	6	1	4	1	2	3	0	2	1	17	5	0	4	0	1	1	1	1	1
IV–2	.94	R	10.3	11.7	0	0	0	0	8	35	1	3	0	0	7	0	1	0	14	12	0	13	1	1	1	0	1	2
		TS	9.7	11.3	1	1	0	0	6	32	0	5	0	1	2	0	2	0	4	11	1	12	1	1	1	0	0	0

TABLE 4 (*Continued*)

Interv. No.	PCI	Form	Direct. rating A.J.	Direct. rating D.S.	Codings corresponding to check list categories																						
					1a	1b	1u	2a	2b	2c	2u	3a	3b	3c	3d	3e	3f	3u	4	5a	5b	5c	5d	5e	5f	5u	OU	I
IV–4	.94	R	4.3	7.2	0	0	1	1	10	11	0	10	9	5	1	4	2	1	11	2	2	1	0	2	1	0	2	1
		TS	6.3	7.9	0	0	0	1	12	16	2	7	9	5	1	1	1	1	16	2	0	0	0	0	1	0	4	0
IV–6	.95	R	2.0	4.4	0	1	0	1	3	2	0	7	15	12	1	0	0	1	1	0	0	0	0	0	0	0	1	2
		TS	2.7	4.8	0	1	0	1	0	2	1	4	16	18	0	0	0	0	1	0	0	0	0	0	0	0	1	2
IV–8	.90	R	1.0	2.8	0	1	0	0	1	0	0	3	10	2	0	0	0	0	0	0	0	0	0	0	0	0	1	2
		TS	2.4	5.0	0	1	0	0	1	2	0	3	8	5	0	0	1	0	1	0	0	0	0	0	0	0	1	1
V–6	.91	R	9.7	9.6	1	1	0	0	17	36	0	7	4	2	5	4	15	0	10	7	1	3	1	1	1	0	11	0
		TS	9.0	9.2	0	1	0	0	24	45	2	7	5	2	2	2	4	2	18	8	0	5	1	4	1	0	4	1
VI–5	.82	R	4.5	6.6	0	1	0	1	3	5	0	7	12	6	0	0	1	0	8	0	0	0	1	0	0	0	0	0
		TS	6.7	5.3	1	1	1	0	8	3	0	5	8	12	0	0	0	0	7	1	1	0	0	1	0	0	0	0
VII–1	.93	R	2.7	4.2	0	5	2	0	5	3	0	7	17	3	3	0	0	1	1	1	0	0	2	0	0	0	0	0
		TS	1.3	3.8	1	6	1	4	4	1	1	9	16	6	0	0	0	1	2	1	0	0	2	1	0	0	0	0
VIII–1	.92	R	4.0	5.7	2	7	0	2	8	3	0	5	10	13	1	0	0	0	3	1	0	0	2	1	0	0	2	4
		TS	6.3	5.8	2	4	0	1	8	7	0	8	12	12	1	0	2	0	2	0	0	1	1	0	0	0	1	2
IX–4	.75	R	10.3	10.6	2	0	0	1	18	39	0	6	0	0	4	2	1	1	24	13	0	6	1	0	1	0	3	4
		TS	11.0	11.2	1	0	0	1	14	33	1	0	0	0	1	2	0	0	16	12	0	5	0	0	0	0	1	3
X–6	.96	R	10.7	10.6	1	0	1	0	14	30	0	7	0	0	3	2	4	0	16	13	1	5	0	1	2	0	2	1
		TS	10.7	11.2	0	0	0	1	13	42	1	2	0	0	2	0	6	0	18	9	1	7	0	0	0	0	2	0

its series by an Arabic numeral; (2) the Profile Comparability Index between the averaged recorded form (R) and the averaged typescript form (TS); (3) the average of the judges' directiveness ratings (AJ) for each form and the derived scores (DS) of directiveness obtained by applying the weights found in Table 1 to the average frequency of use of each category; and (4) the average frequency with which each category of the check list was used to identify procedures in each form. It is from these last figures that profiles may be constructed by plotting the check list category code on the abscissa and the frequencies on the ordinate of a graph.

Analysis of the figures in Table 4 suggests that two major and characteristically different patterns of procedures were identified by the judges. By taking the liberty of paraphrasing the categories of the check list, they may be described as follows: One pattern of procedures seemed to be characterized by a very large proportion of questions. Direct questions (2c) were in greater number than indirect questions (2a, 2b). Several times during the interviews the counselor identified for the client a source or sources of difficulty (3d), that is, the counselor told the client what his trouble was. About half as often as the counselor asked questions, he offered to the client some type of explanation or discussed some aspect of the situation, in the sense of giving information about the situation (4). Then, in roughly the same proportion as problems had been identified for the client (3d) the counselor suggested some activity for the client to carry out (5a). Accompanying this, the counselor often presented or marshaled evidence in favor of one line of action or opposed to another line of action (5c). Very briefly, this pattern presented the picture of diagnosis, explanation, suggested solutions, and persuasion.

A second pattern was characterized by very few questions in an interview, with the more indirect questions (2a, 2b) being more numerous than the direct. Of greater incidence than questions were responses by the counselor which may be characterized as clarifying remarks, that is, remarks which tended to rephrase what the client had said (3a), or to rephrase what the client had not necessarily stated but had expressed

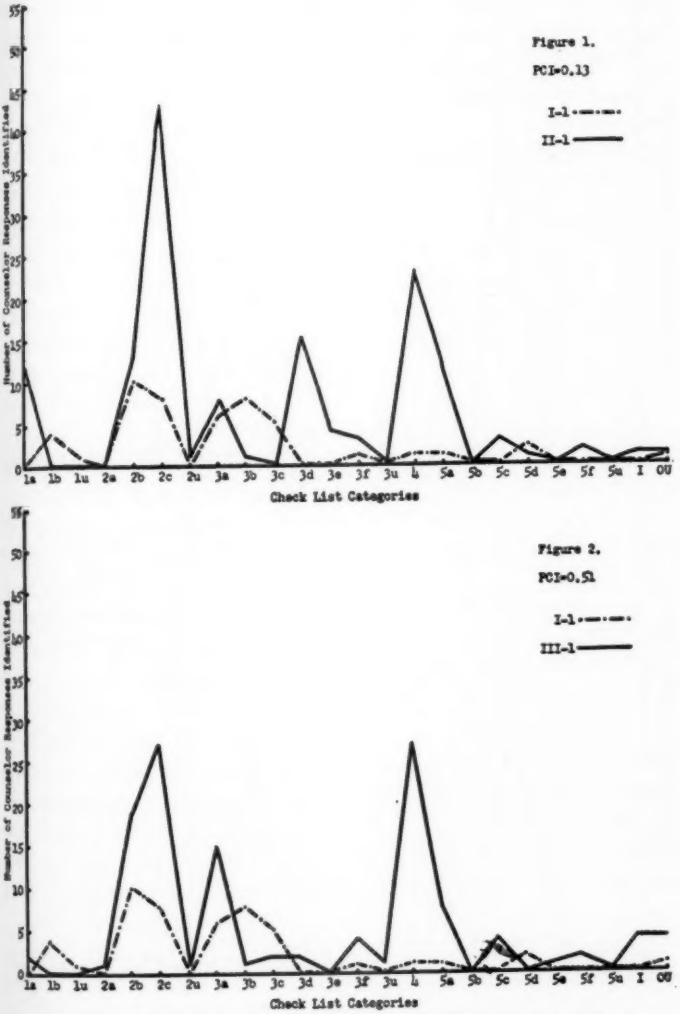
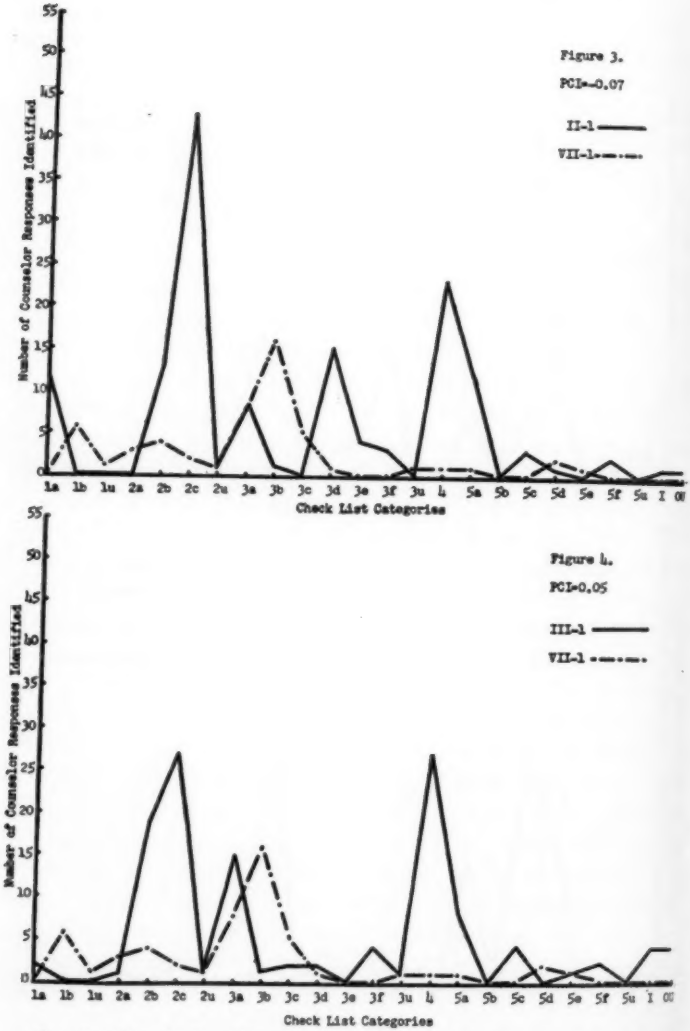


Figure 1.
PCI=0.13
I-1
II-1
Number of Counselor Responses Identified
Check List Categories
1a 1b 1u 2a 2b 2c 2u 3a 3b 3c 3d 3e 3f 3u 4 5a 5b 5c 5d 5e 5f 5u I OU
Figure 2.
PCI=0.51
I-1
III-1
Number of Counselor Responses Identified
Check List Categories
1a 1b 1u 2a 2b 2c 2u 3a 3b 3c 3d 3e 3f 3u 4 5a 5b 5c 5d 5e 5f 5u I OU

FIGS. 1 and 2. Profiles of interviews by counselors differing in point of view.

through his general demeanor (3b). In very few cases were problems ever pointed out to the client (3d) nor was very much information given in discussion (4). Similarly, very seldom



FIGS. 3 and 4. Profiles of interviews by counselors differing in point of view.

did the counselors of this second point of view suggest what the client might do to overcome his difficulties (5a). And even more seldom was there an attempt to marshal evidence in favor

or in disfavor of any line of action (5c). This second pattern seemed to be restricted primarily to such procedures as would help the client to see what his own feelings were, to accept those feelings, and to decide for himself what course of action was to be taken.

Counselors II and III agreed that they accepted more directive principles of counseling while counselors I and VII favored a nondirective approach. In Figures 1, 2, 3, and 4 profiles are presented which compare first interviews conducted by these counselors. These figures indicate contrasts between interviews based on differing philosophies.[11] The profiles presented are composite profiles, that is, averages of typescript and recording profiles. The Profile Comparability Indices are presented with the figures.

In Figures 5 and 6 comparisons are made between counselors II and III, and counselors I and VII. These figures contain the same profiles as appeared in the first four figures but are so rearranged as to compare counselors of like viewpoint. While the PCI values accompanying Figures 1 through 4 suggest that the pattern of procedures used by counselors of differing viewpoint are quite unlike, the PCI values accompanying Figures 5 and 6 suggest that the pattern of procedures used by counselors of like viewpoint are closely alike.

In Figures 7 and 8 the early, middle, and late interviews of counselors I and II, respectively, are presented.[12] Inspection of these figures suggests that a counselor may be quite consistent in the procedures he uses during successive interviews with the same client. It is interesting to note some of the progressive changes which take place during the course of the interviews. In Series I, Figure 7, attempts to define the interview (1a and 1b) occurred predominantly in the first interview, not at all in the last interview; leading (2b and 2c) steadily decreased; and interpretations (3c) tended to increase. In Series II, Figure 8, defining the interview situation (1a) decreased

[11] All of the counselors studied would have agreed as to which viewpoint they favored. For the sake of brevity only the four counselors listed are graphically compared.

[12] While similar data are available for counselor III, they are not presented here because the interviews varied so greatly in length of time.

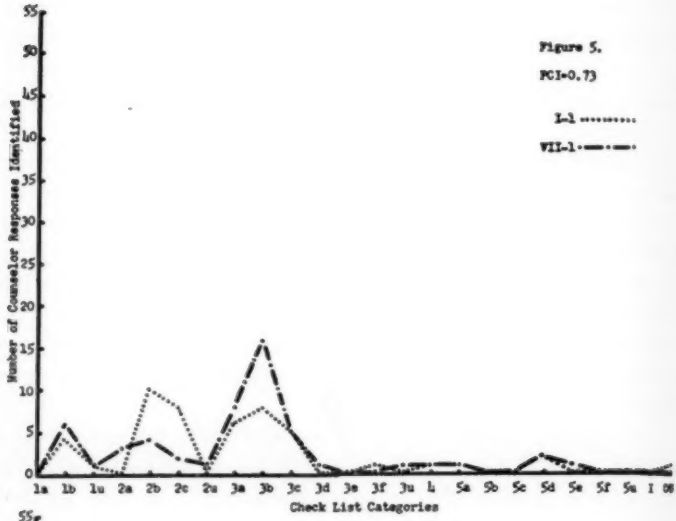


Figure 5.
PCI=0.73
I-1
VII-1
Number of Counselor Responses Identified
Check List Categories
1a 1b 1u 2a 2b 2c 2u 3a 3b 3c 3d 3e 3f 3u 4 5a 5b 5c 5d 5e 5f 5u I 0

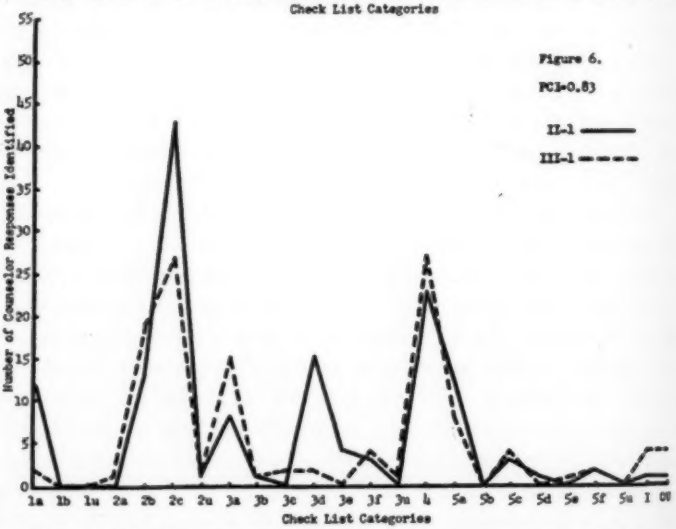


Figure 6.
PCI=0.83
II-1
III-1
Number of Counselor Responses Identified
Check List Categories
1a 1b 1u 2a 2b 2c 2u 3a 3b 3c 3d 3e 3f 3u 4 5a 5b 5c 5d 5e 5f 5u I 0

FIGS. 5 and 6. Profiles of interviews by counselors agreeing in point of view.

from a high frequency to no attempts; leading (2b and 2c) increased as the problem was developed and then decreased markedly at the end; identification of client problems (3d)

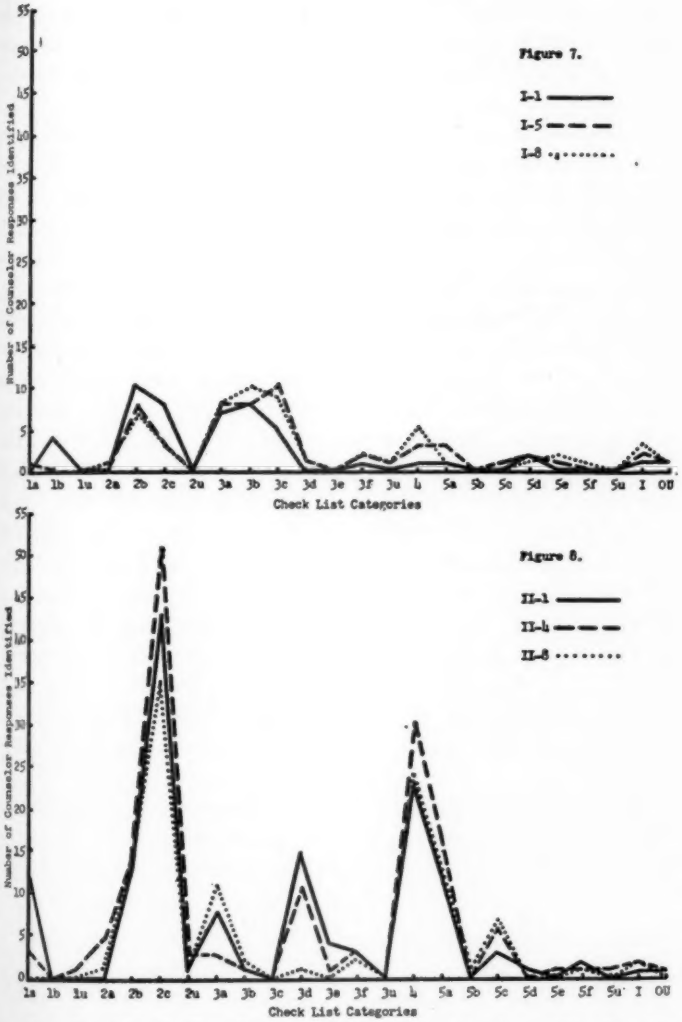


Figure 7.
I-1
I-5
I-8
Number of Counselor Responses Identified
Check List Categories
1a 1b 1u 2a 2b 2c 2u 3a 3b 3c 3d 3e 3f 3u 4 5a 5b 5c 5d 5e 5f 5u I OU
Figure 8.
II-1
II-4
II-8
Number of Counselor Responses Identified
Check List Categories
1a 1b 1u 2a 2b 2c 2u 3a 3b 3c 3d 3e 3f 3u 4 5a 5b 5c 5d 5e 5f 5u I OU

FIGS. 7 and 8. Profiles of early, middle, and late interviews within the same interview series.

decreased markedly; test interpretation (3e) also dropped out; proposals of client activity (5a) increased somewhat; and persuasion (5c) increased as the interview progressed. All of

these changes would probably have been expected upon the basis of most counselors' experiences. That these changes were reflected through the use of the check list suggests that the check list was used with validity.

In the past there has been a great deal of emphasis upon the notion that interviews are primarily unique and in large part influenced by certain unique qualitative factors in the counselors' personalities. Some have felt that these qualitative factors are relatively insensitive to training. It is interesting to analyze Figure 9 in the light of these concepts. Figure 9 presents profiles of the second, fourth, sixth, and eighth interviews held by a counselor-in-training who during the course of the interviews was undergoing instruction which emphasized a non-directive approach. The many marked changes in the pattern of procedures used are evident. The material suggests that the factors within the individual which determine the procedures he is to use are readily subject to training and are not necessarily as unique as is often implied.

The measures presented thus far have related directly to the identification of procedures used by counselors. There are, however, other measures obtainable which may be directly related to the procedures used and are significant for understanding of the nature of counseling processes. One such measure is the ratio of the number of words spoken by the counselor to the number of words spoken by the client (c/c ratio).

If for no other reason, this type of measure is of significance because it constitutes a quantitative measure of the structure of the social interaction within the interview. Table 5 pre-

TABLE 5

Ratios of Number of Words Spoken in Interview by Counselor to Number of Words Spoken by Client

Interview	c/c ratio	Interview	c/c ratio	Interview	c/c ratio
I–1	0.48	III–1	2.94	V–6	0.37
I–5	0.17	III–4	3.95	VI–5	1.46
I–8	0.40	III–7	2.04	VII–1	0.31
II–1	2.47	IV–2	1.30	VIII–1	1.27
II–4	4.02	IV–4	0.70	IX–4	3.25
II–8	2.87	IV–6	0.38	X–6	3.39
		IV–8	0.15		

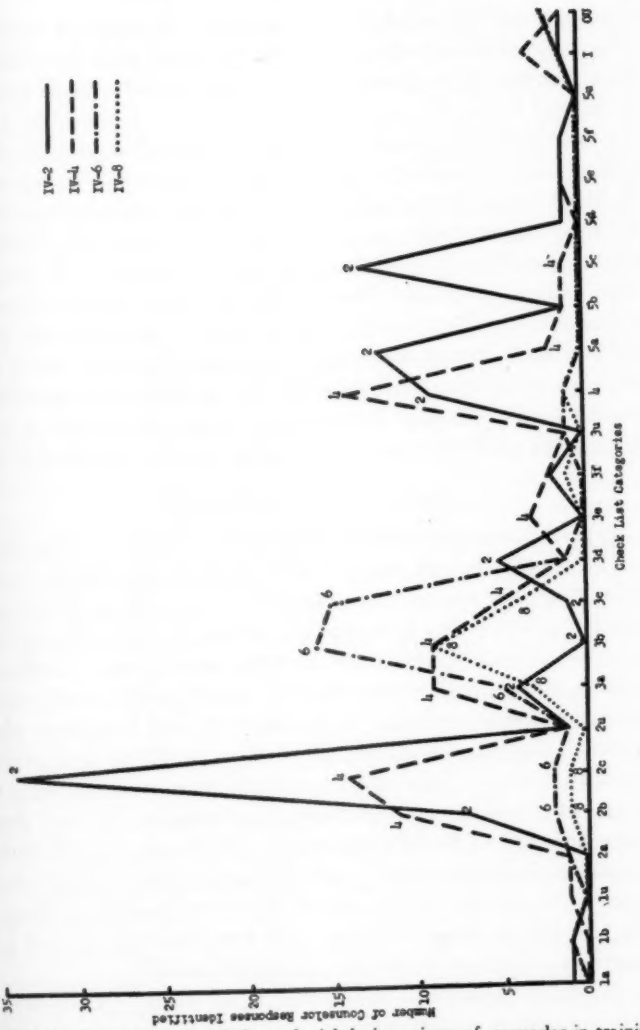


FIG. 9. Second, fourth, sixth, and eighth interviews of counselor-in-training during course of training in non-directive interviewing procedures.

sents these ratios for each interview. The word count was made from the typescript of the interview in each case.

It will be observed in Table 5 that the lowest c/c ratio was

0.15 and the highest 4.02. The highest is some 27 times as great as the lowest. It will also be observed that interviewers tended to be relatively consistent in the proportionate amount of talking from one interview to the next. Counselors I, IV (in interviews 4, 6, and 8), V, and VII tended to talk about half as much as did the clients they interviewed. Counselors II, III, IX, and X tended to talk about three times as much as did their clients. In proportion, the latter group of counselors talked about seven and one-half times as much as did the former group. Further, higher directiveness-nondirectiveness scores were assigned to the latter group. Between the averaged judge-assigned directiveness-nondirectiveness scores and the c/c ratios for each interview the coefficient of correlation equalled 0.94; between directiveness scores derived from averaged profiles and c/c ratios the coefficient also equalled 0.94.

Summary and Implications

The research instrument developed, a check list suitable for use in identifying the number and kinds of procedures used in counseling interviews, was evolved through a theoretical analysis of publications on counseling precepts and an empirical analysis of phonographically recorded interviews. Briefly described, the check list provided for classification of procedures used in carrying out four major interviewing functions which seemed to appear in almost all of the interviews and publications analyzed. These four functions were: (1) defining the interview situation for the client, (2) bringing out and developing the client's problem situation, (3) developing the client's insight and understanding, and (4) sponsoring activities to be carried out by the client and/or fostering decision-making on the part of the client. Under each function were several subheadings, making twenty-four categories in all.

Nineteen interviews were selected for study. They were of about equal duration, reasonably audible, representative of differing points of view on counseling, representative of early, middle, and late interview contacts, and sampled differing degrees of experience among counselors. The thirteen judges who used the check list in identifying and classifying the coun-

selors' procedures were persons who would be accepted as competent counselors. Both range of experience and differences in point of view on counseling were taken into account in selecting the judges.

The judges were familiarized with the use of the check list through written instructions and by conferences with the writer after classifying the counselor responses found in typescripts of interviews. Each judge worked with four typescripts and four phonographic recordings. When working with the typescripts the judges indicated their classification of a procedure by entering a given symbol on the left margin of the page. When listening to the recordings a tally was made on a mimeographed copy of the check list in the appropriate space provided.

Evidence of reliability in the use of the check list came from four sources. The total number of times judges agreed exactly in their identifications was some 4500 times greater than would be expected by chance. Comparisons of general agreement as well as exact agreement also indicated a high degree of agreement. Correlations between the total number of identifications made per interview per judge indicated exceptionally high agreement. Correlations between the patterns of procedures identified were also very high.

Validity was inferred to exist from the manner in which the check list was formulated and employed, from the fact that the check list through weights assigned to its categories ranked the nineteen interviews as to their directiveness in essentially the same order as did the independent judges, and from various comparisons between interview profiles.

The comparisons, although so limited in number as to make generalizing improper, suggested several hypotheses: (1) different viewpoints on counseling are reflected in the patterns of procedures used; (2) similar viewpoints on counseling are associated with similar patterns of procedures; (3) a given counselor is quite consistent in the pattern of procedures he uses throughout a series of interviews with a client; (4) a counselor will tend to adhere more closely to a pattern of procedures consistent with his viewpoint than he will tend to vary his pro-

cedures; and (5) the performance of the counselor may be markedly affected by training.

In the writer's opinion this study has significance through establishing that judges can agree with a high degree of reliability as to the procedures which a counselor follows and through the hypotheses which the study has suggested. The implication also stands out that if quality differences in interviewing can be quantified similar quantification should be possible in other fields of applied social interaction such as salesmanship, social work interviewing, and teaching. Opportunity for quantification also permits diagnostic and achievement measurement.[13]

The next major step to be made is the analysis of the reactions of the interviewee in the hope that cause and effect within the social interaction of the interview may be more adequately understood than at present.

[13] In the construction of material for use in civil service examinations the writer has attempted a job analysis of social work positions based in part upon the interviewing functions incorporated in the check list.

The project was instituted by the California State Department of Social Welfare and carried out under the writer's immediate direction with a grant from the Columbia Foundation. The test, assembled into two parts requiring three to four hours each for completion, included subtests of ability to: (1) read public assistance laws and relate them to practical case material, (2) complete and correct a typical form used in public assistance agencies, (3) discriminate between well and poorly formulated case record material, (4) discriminate between the tone of desirably and undesirably written letters, (5) identify examples of different interview procedures, (6) discriminate between less and more apt responses to clients, (7) identify the significance of the immediate response of a client, (8) identify reasonable diagnostic inferences, and (9) identify various stages and phases in interview progress. Analyses of results obtained by administration of the test to over one hundred caseworkers in rural and urban California communities are as yet incomplete. Analyses are to be made in terms of the use of the material for achievement and aptitude measures.

A COMPARISON OF INTEREST MEASUREMENT BY THE KUDER PREFERENCE RECORD AND THE STRONG VOCATIONAL INTEREST BLANKS FOR MEN AND WOMEN*

J. R. WITTENBORN
Yale University
FRANCES ORALIND TRIGGS
Social Security Board
DANIEL D. FEDER
Illinois State Civil Service Commission†

AMONG the most popular interest inventories used in vocational guidance are the *Vocational Interest Blank* by E. K. Strong, Jr., and the *Preference Record* by G. F. Kuder. They have the common goal of determining relative strengths of interest of any individual in various occupational areas. Scores on the *Vocational Interest Blank* indicate to what extent an individual's pattern of inventory responses approximates the response pattern typical of persons successfully engaged in various occupations. The *Preference Record,* in contrast, furnishes scores which indicate the strength of college students' preferences in broad areas of general interest. The items to be scored for the respective areas of the *Preference Record* were selected on a logical basis and refined by means of item analysis.

Since many of the occupations on which the Strong test has been validated are thought to be logically associated with interests in one or more of the areas of the *Preference Record,* it is possible to establish the common agreement of the two tests by determining what pattern of occupational interests indicated by the Strong *Vocational Interest Blank* is most characteristic of persons scoring unusually high or unusually low in

* The authors wish to express their appreciation to those who have cooperated in this study, and especially to Drs. A. E. Traxler, Janet Weston, and Theodore Sarbin, who have given valuable help and criticism.

† On leave in the U. S. Naval Reserve.

the various areas of preference or interests measured by the *Preference Record.* The present study has attempted to determine the extent of agreement between interests indicated by the two tests.

Description of Tests

There are two forms of the Strong *Vocational Interest Blank.* The women's form contains 410 items which may be scored for 15 different vocations or keys; the men's form contains 400 items which may be scored for 34 vocations. The degree to which an item discriminates one occupation from others determines the scoring weight it will receive when the test is scored for that specific occupation. The individual's raw score for a given occupation is converted into a rating of C, C +, B —, B, B +, or A. A profile is generally prepared on the basis of the letter score for each key, thus showing the relative standings of an individual's interests in each of the occupational areas for which scoring keys are available. A factor analysis of the scores on the 34 vocations for which the men's inventory is scored resulted in a grouping of these vocations into eleven occupational areas. The same type of analysis on the women's inventory resulted in a grouping into five occupational areas.

The inventory is interpreted primarily according to the pattern of interests in an occupational area, and only secondarily according to scores on individual keys. Three types of patterns may be found: primary, secondary, and tertiary, so called according to the intensity of an individual's interests as shown on the profile. A pattern is said to be primary when a preponderance of A and B + scores is found in any one area; any pattern below a primary pattern which receives half or more B or B + ratings is known as a secondary pattern; and a tertiary pattern is any pattern below a secondary pattern receiving at least half of the scores above a C.

Form A of the Kuder *Preference Record* contains 330 items which may be classified as sampling interests in the areas of Science, Computation, Art, Literature, Music, Social Service, and Persuasive activities. Thus, this test provides scales for seven occupational areas: Scientific, Computational, Musical, Artistic, Literary, Social Service, and Persuasive. If an item contributes

to an interest area, it is given a positive score of one point. Otherwise the effect is neutral. The raw scores are converted to percentile ranks based on scores of 800 college students (400 of each sex). It is conventional to consider areas with a percentile rank of 70 or higher as significant although, as in the Strong test, the relative strength of the various interests and the patterns described by them may be highly significant.

There is an important difference between the form of the *Preference Record* items and the items of the *Vocational Interest Blank*. In the latter, the subject is required to indicate for most of the items (which may involve an occupation, a study, an amusement, and so forth) whether he likes, dislikes, or is indifferent toward it; in the *Preference Record*, each item requires that the subject indicate which of two possible situations or courses of action he prefers. The justification for use of the preference type of item has been presented in an article by Kuder.[1]

Description of Data

The data for this study were taken from two sources: the files of the Personnel Bureau of the University of Illinois and the files of the Testing Bureau of the University of Minnesota.

Illinois data consist of Strong *Vocational Interest Blank* and Kuder *Preference Record* profiles for subjects who had completed both of the inventories and who had percentile ranks above the seventy-fifth or below the twenty-fifth centiles on at least one of the keys of the Kuder *Preference Record*. The Minnesota data are comparable to the Illinois data except in two ways: first, the *Preference Record* profiles of all students having percentile ranks above the seventieth centile (as against the seventy-fifth for Illinois data) were analyzed; second, no analysis was made of students whose blanks indicated lack of interest as indicated by scores below the twenty-fifth percentile on the various keys of the Kuder *Preference Record*.

Both studies were well under way before any plan was made for a combined report.

---

[1] Kuder, G. F. "The Stability of Preference Items." *Journal of Social Psychology*, XIX (1939), 41–50.

Method of Analyzing the Data

The procedure for analyzing the data will now be described. First the Kuder *Preference Record* profiles were separated by sex of the individual taking the test because the interests of men and women are measured by different inventories on the Strong *Vocational Interest Blank*. Second, all *Preference Record* profiles on the Scientific Scale which had scores above the seventy-fifth percentile value (for Illinois data, but the seventieth for the Minnesota data) were sorted out and the Strong inventory letter scores for each Strong key were tabulated,[2] thus making a summary profile. The Illinois and Minnesota data were treated separately. For the Illinois data, the *Preference Record* profiles, which had a percentile rank below the twenty-fifth centile on this key, were also sorted out and the Strong inventory letter scores were tabulated. Third, the *Preference Record* profiles were put back into a group and redistributed in the same manner for the Computational Scale on the Kuder *Preference Record*. The whole process was then repeated for this scale and for each of the five remaining scales for each sex. Three tabulations for each Kuder *Preference Record* scale of the letter scores on the Strong *Vocational Interest Blank* for each sex resulted. For the purpose of direct comparison of these data, a summary profile of the tabulated Strong letter grades for each scale on the *Preference Record* for each sex was prepared in a blank Strong profile. The method for making these profiles was as follows: the median score on each Strong key was found by taking the letter value containing the midpoint for the tabulation on each key. Where there were an even number of points and thus were no exact midpoints, the higher of the two center points was taken. Thus fourteen summary profiles resulted, one for each of the seven scales on the *Preference Record* for each sex. Each profile has on it:

1. The median score for each Strong key for the Illinois students receiving scores on the Kuder *Preference Record* above the seventy-fifth percentile.

[2] It is possible by this criterion for a Kuder *Preference Record* profile to have a score on more than one key fall within the area to be studied. In such a case the profile was included in as many distributions as there were scores within the range being studied.

2. The median score for each Strong key for the Minnesota students receiving scores on the Kuder *Preference Record* above the seventieth percentile.

3. The median score for each Strong key for the Illinois students receiving scores on the Kuder *Preference Record* below the twenty-fifth percentile.

On the basis of these summary profiles it is possible to determine the extent of agreement between interests indicated by the two tests.

Analysis of Data

It has already been noted that these two tests have a common goal of determining relative strengths of interests of any individual in various occupational areas. Thus some common relationships might be expected. For instance, students having significantly strong interests as measured by the Kuder *Preference Record* Scientific Scale might be expected to have a primary or secondary pattern of interests on either Group I or Group II (or both) of the Strong *Vocational Interest Inventory* profile.

Throughout this discussion the actual distributions will be compared to the "to be expected" distribution to note the extent of agreement.

One fact should be noted about both the men's and women's profiles, however. Those persons used to interpreting the Strong profiles are aware that a pattern in Groups IV and VIII (and to some extent IX) on the men's blank and Group II on the women's blank appears with much more frequency than do patterns on the other Groups. Thus the appearance of these patterns must not be interpreted to be as significant as the appearance of other patterns. (Group IV on the men's key is often referred to as descriptive of interests of "men in general" and Group II on the women's key as descriptive of "women in general".)

The part which is being played by the Illinois low data should be noted throughout the interpretation. They not only help to indicate the strengths of certain stated relationships; from these data also appear certain very interesting negative relationships which from time to time will be mentioned.

While they sometimes defy interpretation, the presence of these negative relationships is challenging for further study.

The profiles will now be analyzed separately in the order in which they appear.

Profile 1 in Figure I is the tabulation of the letter scores on

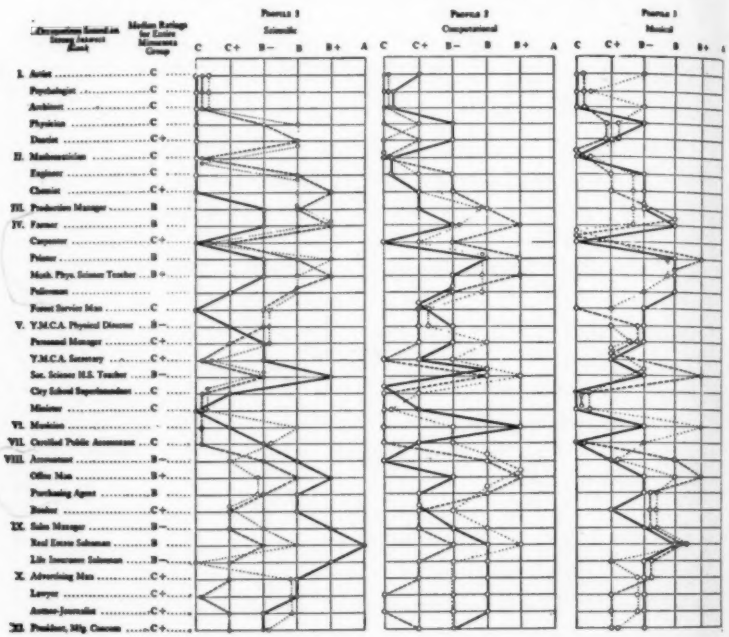


FIGURE I (*Continued on next page*)
Median Profiles on the Strong *Interest Blank* of High and Low Groups of Men on the Kuder *Preference Record*

The solid line refers to the low Illinois group, the broken line to the high Minnesota group, the dotted line to the high Illinois group.

the Strong keys when tabulated for the Scientific Scale on the Kuder *Preference Record.* From Kuder's description it might be expected that students scoring high on this scale would have strong patterns on Groups I and II of the Strong *Vocational Interest Blank.* Kuder describes the key as follows:

Scientific activities: The following activities are representative of those which are scored for this scale when preferred to activities which are of a less experimental nature: use a microscope, work me-

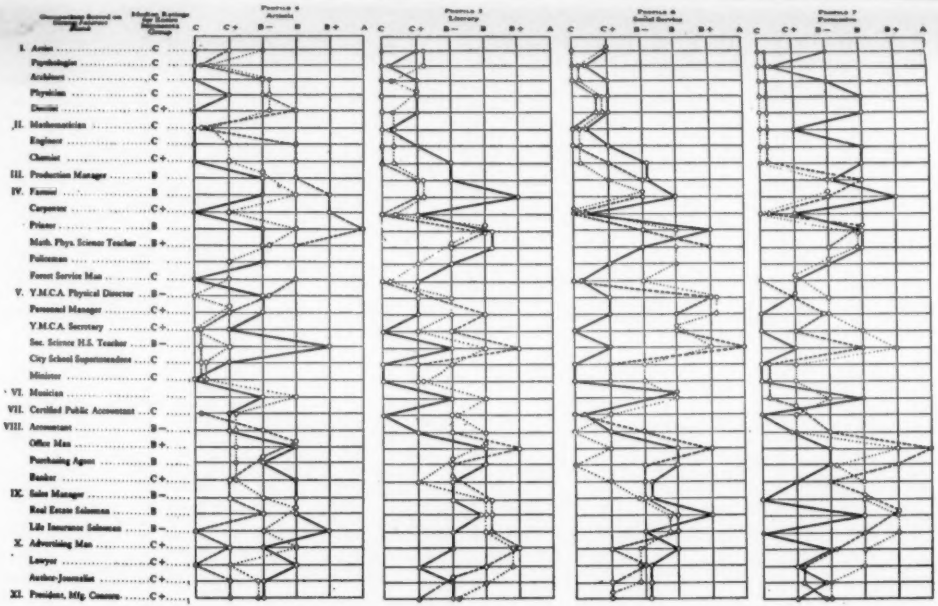


FIGURE I (*Continued*)

Median Profiles on the Strong *Interest Blank* of High and Low Groups of Men on the Kuder *Preference Record*

The solid line refers to the low Illinois group, the broken line to the high Minnesota group, the dotted line to the high Illinois group.

chanical puzzles, perform laboratory experiments, read about the causes of various diseases, visit an exhibit of scientific apparatus, experiment with growing flowers in different soils, try out variations of a recipe for candy mints to get a better mint, take a vacuum cleaner apart to see what is wrong with it, develop new and better varieties of fruit, belong to an amateur astronomy club, develop a cure for infantile paralysis.

Occupations suggested for consideration by students with high scores in this field: research scientist, chemist, physicist, inventor, geologist, criminologist, psychologist, biologist, electrical engineer, laboratory technician, optician, physician, surgeon, pharmacist, bacteriologist, oculist, science teacher, explorer, archeologist.[3]

An examination of Groups I and II on the summary profile reveals that a tertiary pattern is found for both the Illinois high data and the Minnesota data on Group I and a secondary pattern on Group II. Thus it seems that the Scientific Scale on the Kuder inventory is more closely related to types of interests described by Group II of the Strong inventory than Group I. However, it should be noted that while no significant scores are found for three of the keys in Group I, Artist, Psychologist, and Architect, in both Illinois and Minnesota high data, Dentist and Physician are positively related.

The Illinois data which indicate lack of interest on the *Preference Record* also indicate no interest in these two keys on the Strong. It seems that this same group has much stronger interests in contact business and literary pursuits.

In summary, it might be said that given a Kuder *Preference Record* profile with a percentile rank of seventy-five or over on the Scientific Scale, it is probable that were the same student given the Strong, his scores might be predicted to form a pattern in Group II and there might be an indication of interest on the Physician and Dentist keys of the Strong. It would also be predicted that no significant support would be given this pattern from the first three keys in Group I. Kuder, in his description of the scientific key, thought that psychologists would be included by this scale. He does not include artist or architect.

In Profile 1, as in many of the other profiles, certain negative relationships are apparent. For example, the Illinois high

[3] These descriptions are from the manual for this edition of the Kuder *Preference Record.*

group and the Minnesota data on the Kuder Scientific Scale show ratings of "C–" in the Strong Life Insurance Salesman Key, but the low Illinois group has a rating on this occupation of "B +." Such negative relations have not yet been explored, but it seems plausible to suppose that, when better understood, they may have considerable significance for the vocational counselor.

Profile 2 in Figure I is a summary for the Computational Scale on the *Preference Record.* Kuder describes this scale as follows:

> Activities involving computation: Among the activities included in this scale are: be treasurer of a club, study algebra, study geometry, be the cashier in a restaurant, collect figures on traffic accidents, prepare tables of figures on public health, keep the official statistics for a football game, keep a record of the traffic past a certain point, determine the cost of producing a newly-developed dishwasher, read about a new device for making quick calculations.
>
> Occupations suggested for consideration: certified public accountant, actuary, statistician, arithmetic teacher, physicist, bookkeeper, credit manager, mathematician, bank teller, broker, purchasing agent, insurance salesman, bond salesman, cost accountant, surveyor, auditor, architect, billing clerk.

From this description it would be expected that there might be a significant pattern on Groups VII and VIII and some relationship to the mathematician interests on the Strong. The pattern for Group VIII is a secondary one. As discussed earlier, this pattern is quite common so that it is somewhat difficult to say that its presence here indicates a meaningful relationship, although the low students on the computational scale are distinctly lower in this group of Strong scores. The expected relationship between this Kuder scale and the mathematician key certainly indicates that the two are not measuring the same type of interest. Neither is a relationship indicated between the Certified Public Accountant Strong key and this Kuder scale.

Profile 3 depicts the Strong tabulation for the Kuder Musical Scale for Illinois data only. Kuder's description of this scale indicates that a pattern should be expected for the Strong Musician key only. The most which can be said for this tabulation is that the strongest point for the Illinois high tabulation

is for this group and that the Illinois low tabulation is two letter points lower. Whether or not the Kuder and Strong keys are measuring the same interest is still in question.

Profile 4—From Kuder's description of his Artist Scale, it is difficult to expect any pattern. The results can probably best be interpreted by saying that the tabulation indicating no interest is in closest agreement with what might be expected. Whether the C + and B – on the Artist key and the two B –'s on the Architect key are significant is questionable. Actually the strongest pattern found is in Group IV for the Minnesota students although this result is not borne out in terms of differences within the Illinois group.

Profile 5—Kuder describes the Literary Scale as follows:

> Literary activities: This scale includes the following activities, among others: write a book review, write the greetings for greeting cards, do crossword puzzles, enter a short-story-writing contest, browse in a library, collect first editions of books, write an advertisement, take a course in modern literature, direct a play, be editor of a college yearbook, write the script for a radio program, write a clever one-act play, read about how language is changing.
>
> Suggested occupations: author, journalist, editor, novelist, poet, literary critic, playwright, dramatic critic, actor, actress, teacher, advertising copy writer, manuscript reader, proofreader, secretary, librarian, reporter.

About the only pattern to be expected from this description would be in Group X. A secondary pattern is found for both Illinois and Minnesota high data. A tertiary pattern is found for Illinois low data. It is interesting to note that there is no significant difference between the three sets of data on the Author-Journalist key. The very strong indication in Group V that the Teacher of Social Science key is related to this Kuder scale should be noted.

In Profile 5, there is another interesting negative relation for the occupation of Farmer. Here the high Illinois data and the Minnesota data each have a median Strong rating of "C," whereas the low Illinois data have a median Strong rating of "B +." Although the number of cases in the high Illinois data is rather small, only eleven, there are 24 cases in the Minnesota data and 40 cases in the low Illinois data. It does not seem likely that this relationship is accidental.

Profile 6—Kuder's description of the Social Service Scale is as follows:

Social service activities: The term "social service" is used here in a broad sense to include activities indicative of an interest in people with the emphasis on those which involve personally helping people in a fairly direct way. Among the activities in this scale are: help a friend pick out a new hat, get strangers acquainted at a party, teach people how to keep in good health, do social welfare work, read an article on how to treat mental ills, direct a playground for underprivileged children, be a vocational counselor, conduct a group discussion on student problems, spend a vacation as counselor in a summer camp.

Suggested occupations: social service worker, psychiatrist, school physician, vocational counselor, dean, minister, priest, rabbi, missionary, playground director, nurse, scout leader, camp director, kindergarten teacher, case worker for a juvenile court, sociologist, teacher, 4-H Club supervisor, county demonstration agent, personnel manager.

This description would lead one to think that a significant pattern of interests should be found in the Strong Group V. This expectation is borne out. It is given weight when the negative factor contributed by the Illinois low data is studied.

Profile 7—Kuder describes the Persuasive Scale as follows:

Persuasive activities: This scale includes activities in the general field of convincing, persuading, managing and selling, as compared with activities less of this type. Among them are: give a pep talk at a school assembly, sell tickets for a play, get information on a problem by interviewing people, get people to come to a class picnic, solicit advertisements for a college paper, take a course in salesmanship, visit the offices of an advertising agency, be president of the student council of a college, study public speaking, be a guest speaker at a club meeting, direct a play.

Suggested occupations: salesman, business executive, psychiatrist, personnel director, lawyer, diplomat, consul, buyer, hotel manager, recreation director, sales demonstrator, radio announcer, factory foreman, policeman, complaint adjuster, store clerk, retail merchant, efficiency expert, bond salesman, press agent, army officer, theater manager, bill collector, insurance agent, labor leader, house-to-house salesman, advertising manager, sales manager, advertising writer.

All of these activities do not fall into one well-defined Strong Group, but from the description a pattern in at least Group X would be expected. The expected pattern is found and the relationship between the Kuder scale and the Strong Group is present. Kuder mentions personnel manager, and the rela-

tionship between this Kuder scale and that Strong key in Group V is apparent, as it is also for the Social Science Teacher key.

Turning to a consideration of the summary profiles for women, it must be borne in mind that for the Strong there are separate tests for women and men but the Kuder is the same for both sexes. The patterns to be expected on the summary profiles for the Kuder scales for women would be the same as for men. It should be borne in mind that Group II, on the Strong Inventory, appears with much more frequency than does any other one pattern. The number of cases involved in each tabulation should be carefully considered here also, for in some cases it is quite small.

Profile 8 in Figure II depicts the mean Strong profile for the Kuder Scientific Scale. The Illinois high data should be disregarded because of the very small number of cases. The pattern to be expected would fall in Group IV. It will be noted that a secondary pattern does exist in this area for the Minnesota high data (24 cases) and that the significance of this pattern is strengthened by the lack of pattern for Illinois low data (37 cases).

Profile 9 in Figure II is the Strong summary profile for the Kuder Computational Scale for women. There is hardly a pattern to be expected if Group II is to be disregarded. However, in this case perhaps the type of work which falls in Group II is accurately described by the Kuder scale. It should be noted that there is a primary pattern on Group II for both the Illinois and Minnesota high data.

In this case the pattern for the group which lacks interest is perhaps more interesting. This group has a secondary pattern only in Group II, and a secondary pattern in Group I and an "A" score on the Teacher of English key. This would suggest that those people who lack interests as measured by the Computational Scale of the Kuder may have creative or verbal interests.

On Profile 10, the Kuder Musical Scale, there is no pattern to be expected and no clear-cut pattern is evident. The only place where any differentiation is clear is on the Lawyer key of Group IIIa, and the number of cases is so small (8) that the significance of this difference is questionable.

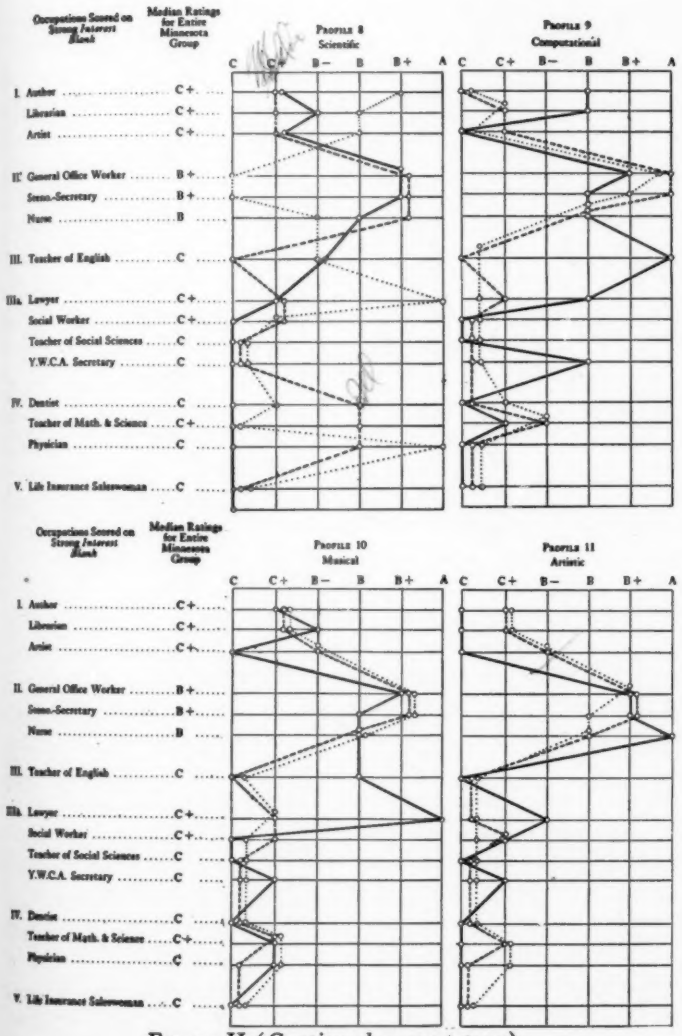


FIGURE II (*Continued on next page*)

Median Profiles on the Strong *Interest Blank* of High and Low Groups of Women on Kuder *Preference Record* Scores.

The solid line refers to the low Illinois group, the broken line to the high Minnesota group, the dotted line to the high Illinois group.

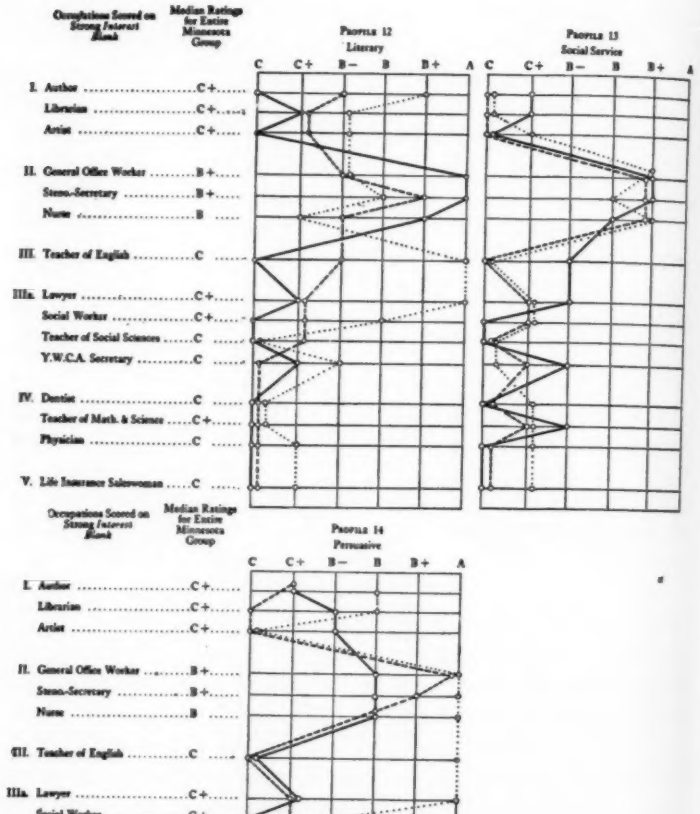

Occupations Scored on Strong Interest Blank
Median Ratings for Entire Minnesota Group
Profile 12 Literary
Profile 13 Social Service
C C+ B− B B+ A
I. Author ..... C+
Librarian ..... C+
Artist ..... C+
II. General Office Worker ..... B+
Steno.-Secretary ..... B+
Nurse ..... B
III. Teacher of English C
IIIa. Lawyer ..... C+
Social Worker ..... C+
Teacher of Social Sciences C
Y.W.C.A. Secretary C
IV. Dentist C
Teacher of Math. & Science ..... C+
Physician C
V. Life Insurance Saleswoman C
Occupations Scored on Strong Interest Blank
Median Ratings for Entire Minnesota Group
Profile 14 Persuasive
C C+ B− B B+ A
I. Author ..... C+
Librarian ..... C+
Artist ..... C+
II. General Office Worker ..... B+
Steno.-Secretary ..... B+
Nurse ..... B
III. Teacher of English C
IIIa. Lawyer ..... C+
Social Worker ..... C+
Teacher of Social Sciences C
Y.W.C.A. Secretary C
IV. Dentist C
Teacher of Math. & Science ..... C+
Physician C
V. Life Insurance Saleswoman C

FIGURE II (*Continued*)

Median Profiles on the Strong *Interest Blank* of High and Low Groups of Women on Kuder *Preference Record* Scores.

The solid line refers to the low Illinois group, the broken line to the high Minnesota group, the dotted line to the high Illinois group.

On Profile 11, the Kuder Artistic Scale, the Illinois and Minnesota high data are in agreement, but the strength of the pattern on Group I, where it would be expected one would appear, is not great enough to give very much evidence that the two tests are measuring the same type of interests in this case. There is some differentiation between the high and low groups, however.

Profile 12 compares the Strong with the Kuder Literary Scale. It would be expected that the strongest pattern of interest would appear on Group I and perhaps III. These patterns do appear and are differentiated from the low group somewhat, but neither the strength of the pattern nor the differentiation is great. It should be noted that the pattern on Group II for the low data is much stronger than for the high data. There is a primary pattern on this group for the low data and only a tertiary for the high data. This would suggest that the group of women having interests in Groups I and II do not have the same type of interests as are depicted by Group II.

On Profile 13, the Kuder Social Service Scale, a pattern might be expected to appear in Group IIIa. Such a pattern is not present, which suggests that the Strong test and the Kuder test are not measuring the same thing in this area. The most interesting suggestions from this pattern are from the low data patterns. However, the number of cases is small (9) and it is therefore questionable whether significance should be attached to it.

For Profile 14, the Kuder Persuasive Scale, there is really no pattern on the Strong profile which might be expected unless it would be V or II. A tertiary pattern does appear for Minnesota high data in Group V but is not supported by Illinois high data (though this may be because of the small number of Illinois high cases). A primary pattern does appear for both Illinois and Minnesota high data, whereas only a secondary pattern is present for Illinois low data, which supports the expected differentiation in this area. The pattern which appears in Group I for Illinois high data and on the Lawyer and Y.W.C.A. Secretary keys probably is due to the small number of cases involved. Logically it could be reasoned that the

latter two might be expected. They lack support, however, and therefore very little significance should be attached to them.

Conclusions

1. There are indications from these data that evidences of lack of interests in certain areas may give important additional clues for vocational counseling. These data may be used to support positive interests and suggest further investigations of some positive interests which might otherwise be overlooked.

2. Because of the frequency with which they appear, there is confirmation for the fact that a pattern of interest in Group II on the women's Strong profile and Groups IV and VIII on the men's should be carefully investigated before receiving the same emphasis as patterns on the other groups receive.

3. The data throw the following light on the comparative interpretation of the Kuder and the Strong interest inventories:

a. The agreement expected between the Scientific Scale of the Kuder *Preference Record* and Group II on the men's Strong and Group IV on the women's Strong is found by this investigation. There is indication that the Physician and Dentist keys in Group I on the Strong also measure to some extent the same type of interests as does the Scientific Scale on the Kuder. However, there is little indication that what is being measured by the Artist, Psychologist, and Architect keys is also being measured by the Scientific Scale of the Kuder *Preference Record.*

b. A comparison of the Computational Scale of the Kuder *Preference Record* and the tabulation of Strong keys indicates positive relationship between Group VIII on the men's Strong and Group II on the women's. There is no indication of close correspondence between this scale on the Kuder *Preference Record* and the Mathematician key or the C.P.A. key on the men's Strong.

c. Little agreement was expected or found between the Kuder Musical Scale and the men's or women's Strong.

d. There is some indication of agreement between the Artistic Scale on the Kuder *Preference Record* and the Artist and

Architect keys on the men's Strong but very little indication of agreement for any key on the women's Strong.

e. The expected correspondence between the Literary Scale on the Kuder *Preference Record* and the men's Strong is greater than for the women's, though there is some evidence to support both.

f. The extent of agreement between the Social Service key and the corresponding key of the men's Strong is quite marked, but the same extent of agreement fails to appear on the women's Strong.

g. For the Persuasive Scale on the Kuder *Preference Record* we find a pattern corresponding to the "to be expected" pattern on the men's Strong. It was difficult to determine what pattern is to be expected on the women's Strong and actually no clear-cut results were noted here.

These data should be considered tentative. There are now available two additional keys on the Kuder *Preference Record,* Form B. A similar study might be made using this edition of the *Preference Record.* The extent to which this study and the projected study agreed would give valuable additional information.

Addendum

The study of the Kuder *Preference Record* and the Strong *Vocational Interest Inventory* was repeated by Triggs using the 1942 edition of the Kuder *Preference Record.* The purpose of repeating the study was to check on the stability of the patterns obtained from the first study and to obtain like information on the Mechanical and Clerical Scales which were added to the 1942 edition of the Kuder test. The method used was the same as that of the original study. The sampling roughly parallels that of the Illinois sampling in the first study. (N = 166 men, 102 women).

The profiles obtained in the second study that correspond to those of the first study for both sexes are so similar that they will not be reproduced; however, differences will be commented upon. The Mechanical and Clerical Scales were not available on the first form of the test; therefore, the profiles obtained in

the second study for these two scales will be described but will not be reproduced.

The profile for men on the Mechanical Scale of the 1942 Kuder *Preference Record* in general very closely approximates Profiles I and II, the profiles for the Scientific and the Computational Scales. This close relationship is borne out also in the profiles constructed (but not reproduced) for the Scientific and Computational Scales of the 1942 *Preference Record.* The following differences seem worthy of comment, however: (1) The pattern in Group I of the profile on the Strong *Vocational Interest Inventory* for the Mechanical Scale of the Kuder *Preference Record* more nearly approximates that of the Scientific than the Computational. This difference might be interpreted as follows: students having measured mechanical and scientific interests on the Kuder *Preference Record* may also have measured interests on the Architect and Physician keys of the Strong inventory but this evidence is less pronounced for those who have measured computational interests on the Kuder; (2) there is slight evidence to indicate that those students who have measured scientific and computational interests on the Kuder are more likely also to have linguistic interests as measured by the Strong inventory (Group X) than are those who have measured mechanical interests on the Kuder; (3) the pattern of interests in business detail (Group VIII) on the Strong *Vocational Interest Inventory* is more pronounced for those persons having measured interests on the Computational Scale of the Kuder than for those having measured interests on the Mechanical and Scientific Scales.

There were only sixteen women who scored high on the Kuder Mechanical Scale. No positive pattern emerged on the Strong inventory profile except a pattern in Group II which is not as predominant as usual. This profile showed no differences between those scoring high and those scoring low on the Clerical Scale. There is negative evidence which is similar to findings on the data for men. Mechanically inclined women do not generally have linguistic interests, but those who score low on the Mechanical Scale do seem to have linguistic interests.

From the profile on the Strong inventory for those women

scoring high on the Clerical Scale of the 1942 Kuder *Preference Record,* it would seem that these women also have interests like successful general office workers and stenographer-secretaries as measured by the Strong *Vocational Interest Inventory.* This conclusion is interesting in light of the fact that for men a similar relationship does not seem to exist between the business detail keys on the Strong inventory and the Clerical Scale of the Kuder *Preference Record.*

The profiles for women are not so stable from the first to the second study as they were for men, though marked interpretable changes are not noted.

The author feels that information on the Clerical Scales should be especially valuable to counsellors using these tests in vocational counselling, since evidence indicates that counsellors tend to interpret the Kuder Clerical Scale and the business interest patterns on the Strong inventory for men and women as being closely related. These data would not support such interpretation.—F.O.T.

AN EVALUATION OF CERTAIN TESTS IN PREDICTING MECHANIC LEARNER ACHIEVEMENT[1]

ELDON E. JACOBSEN
Private, FC, U. S. Army, Hq. & Hq. Company D. E. M. L.
Personnel Replacement Depot, Pittsburg, California

Introduction

Need for the proper placement of trainees in several aircraft classes was the cause for investigating the validity of different tests as predictors of success in aircraft trades. In a government training program, known as the Mechanic Learner Program, several thousand civilians are being trained for numerous skilled aircraft jobs. The persons trained are later employed at numerous air depots as mechanics.

Utah's training centers, in an effort to place the individual more efficiently in the aircraft trade best suited to his aptitudes, inaugurated a personnel program involving the extensive use of tests. Since no battery of tests had been validated for the prediction of achievement in the Mechanic Learner Program, it was necessary to determine this validity for a selected group of tests designed to predict achievement in related fields. The central problem of the study reported here was to check the validity of certain tests as predictors of the achievement of Mechanic Learners as determined by instructors' ratings.

The Criteria of Achievement

Five Mechanic Learner classes receiving training at the Utah State Agricultural College were selected for study. Samples of eighty students from each class were studied. The classes involved were Aircraft Instrument Repair, Aircraft

[1] This article is a condensation of the writer's Master's thesis of the same title, which was done under the direction of Dr. Arden Frandsen, Department of Psychology, Utah State Agricultural College, Logan, Utah. A copy is on file in the college library.

Engine Overhaul, Aero Repair Mechanics, Machine Shop, and Aircraft Electricity. The classes were composed of male students ranging in age from seventeen to forty-four years. Although women were eligible for training, none were included in this study because of their scarcity in the early months of the training program. The average education of the subjects was high-school graduation, although individual cases varied from attendance in high school to attendance in college.

Mechanic Learner classes. Each of the courses involved approximately 580 hours of class and laboratory work. In an eight-hour day, two hours were spent in lecture and study for technical and related subjects of the craft. Trainees spent the remaining six hours in the various shops doing laboratory work. Here instruction was mainly individual explanation and demonstration. As a part of the 580 hours, each class studied approximately thirty hours of Blueprint Reading, five hours of Safety, and five hours of General Orientation. One class, Machine Shop, included thirty hours of Shop Mathematics, five hours of instruction on Materials and sixty hours of use of Precision Tools.

Method of rating achievement. Within each class, trainee achievement was rated periodically by the instructors. Ratings were marked on the "Mechanic Learner Appraisal Scale." The scale involves ratings on seven separate traits: Learning, Speed and Coordination, Workmanship, Working Relations, Responsibility, Interest and Enthusiasm, and Personal Fitness for Occupation. Each trait is divided into a ten-point scale with five verbal descriptions, ranging from low to high, under the numerical scale, to aid the instructor in making a more qualitative judgment in quantitative values.

Ratings were obtained after the first two weeks of training and each month thereafter until training was completed at the school. During a learner's period of training, three, and many times four, ratings of his achievement were recorded.

One method for obtaining greater reliability was to average these ratings, resulting in one over-all criterion to correlate with the predictive devices, instead of using three or four different periodic ratings. In some cases, final ratings on one

individual involve ratings by two or three instructors, due to substitutions and replacements on the instructional staff. Analysis of a sample of 95 cases from two classes revealed a high relation between each of the traits rated and the total of the ratings, as shown in Table 1. The Total Fitness score only was therefore used as the criterion.

TABLE 1

Correlations between Total Fitness for Occupation and Separate Trait Ratings for Trainees in Aero Repair Mechanics. (N = 95)

	Learning	Speed	Work-manship	Working relations	Respon-sibility	Interest
Total Fitness ..	.90 ± .01	.91 ± .01	.95 ± .01	.84 ± .02	.97 ± .00	.89 ± .01

Predictive Devices

Since the skills required for success in any one class are complex, it would appear futile to try to predict total achievement with any one test measuring a simple trait. For this reason, a large battery was chosen. There is little need to discuss previous studies of validation for most of these tests, because of their availability in previous writings. These tests include: Bennett's *Test of Mechanical Comprehension,*[2] *Revised Minnesota Paper Form Board Test,*[3] Johnson O'Connor's *Finger Dexterity Test* and *Tweezer Dexterity Test,*[4] *Minnesota Manual Dexterity Test* (*Placing and Turning*),[5] Crawford's *Test of Tridimensional Structural Visualization*[6] and a combined score of the *Pressey Senior Classification* and *Pressey Verifying Tests.*[7]

One test, the *Arithmetic Prognosis Test,*[8] will be briefly mentioned. It is made up of a sampling of arithmetic prob-

[2] Bennett, George K. *Test of Mechanical Comprehension.* Form AA. New York: The Psychological Corporation, 1940.

[3] Likert, R. and Quasha, William H. *Revised Minnesota Paper Form Board Test.* Series AA. New York: The Psychological Corporation, 1941.

[4] O'Connor, Johnson. *Finger Dexterity Test* and *Tweezer Dexterity Test.* Hoboken, New Jersey: Stevens Institute of Technology.

[5] Ziegler, W. A. *Minnesota Manual Dexterity Test.* Minneapolis, Minnesota: Educational Test Bureau, Educational Publishers, Inc., 1932.

[6] Crawford, John Edmund. *Test for Tridimensional Structural Visualization.* New York: (Distributor) Psychological Corporation, 1940.

[7] Pressey, S. L. and L. C. Pressey. *Senior Classification Test* and *Pressey Senior Verifying Test.* Bloomington, Illinois: Public School Publishing Company, 1940.

[8] Frandsen, Arden. *Arithmetic Prognosis Test.* (Uupublished) Utah State Agricultural College, Logan, Utah.

lems which represent the fundamental skills of mathematics usually considered essential for the performance of general mechanics. It consists of items of addition, subtraction, multiplication, division, decimals, fractions, problem situations, algebraic signs, and problems with simple geometric figures. The test was arranged locally by Dr. Arden Frandsen for use in the college guidance program. Although no data on it had been published, it had been found to correlate significantly with college freshmen grades in mathematics and chemistry. Early results on the test also showed that it provided a normal distribution of scores. These data were an indication of its probable usefulness in this situation.

Interdependency of predictive devices. In order to discover the extent to which the selected tests overlap, intercorrelations were obtained from test scores made by a sample of 90 students in the Aircraft Instruments classes.

The results shown in Table 2 substantiate the generalization that there is little relationship among the tests. However, some results show that a few of the tests are measuring common factors, at least to a small degree. The correlation of $.39 \pm .06$ between the *Pressey* tests and the *Arithmetic Prognosis Test* reveals the probability that both are measuring, somewhat, the same traits. The correlation is not sufficiently high to conclude that one is a duplication of the other. Their significance as individual predictive devices would be the determining factor in evaluating their multiple use.

There is also an indication of interrelationship between the *Mechanical Comprehension Test* and the *Arithmetic Prognosis Test,* as shown by the correlation of $.24 \pm .07$. The relationship is obviously too small to exclude either test on the basis that it is being duplicated by the other.

Some of the manipulative tests have some degree of correlation, such as, *Finger Dexterity* with *Tridimensional Structural Visualization,* $.22 \pm .07$; and *Tweezer Dexterity* with the *Placing Test,* $.26 \pm .07$. These relationships are somewhat significant statistically, but do not reveal much overlapping of the functions measured.

TABLE 2

Intercorrelations among Predictive Measures for Trainees in Aircraft Instruments. (N = 90) *

	Arithmetic Prognosis	Minnesota Paper Form Board	Finger Dexterity	Tweezer Dexterity	Tridim. Structural Visualiz.	Mechanical Comprehension	Placing Test	Turning Test
Pressey Tests	.39 ± .06	.02 ± .07	.07 ± .07	.16 ± .07	.08 ± .07	.16 ± .08	.16 ± .07	.12 ± .08
Arithmetic Prognosis		.01 ± .07	.08 ± .08	.06 ± .07	.11 ± .07	.24 ± .07	.08 ± .07	.10 ± .08
Minn. Paper Form Board			.14 ± .07	.10 ± .07	.20 ± .07	.00 ± .07	.06 ± .08	.00 ± .08
Finger Dexterity				.17 ± .07	.22 ± .07	.04 ± .08	.21 ± .08	.06 ± .08
Tweezer Dexterity					.11 ± .07	.14 ± .08	.26 ± .07	.20 ± .08
Tri. Struct. Visualiz.						.18 ± .08	.19 ± .08	.11 ± .08
Mechanical Comprehension							.21 ± .09	.14 ± .09
Placing Test								.27 ± .07

* Because it was impossible to administer all tests to every individual, the sample is not always the same. In one instance, the sample was just 64, but in a majority of cases, N = 90.

A further generalization might also be made that where relationship is shown, it occurs between two measures involving abstract thinking, rather than between a measure of abstract thinking and a manipulative test.

In view of the small amount of duplication found in the selected tests, it was decided to retain them all in the battery of tests to be checked against the criterion in each class.

Predictive Significance of Tests in Determining Mechanic Learner Achievement

Correlations between tests and the achievement criteria are given in Table 3. Inasmuch as each class involves, in the main, different skills, each will be considered separately.

TABLE 3

*Correlations between Predictive Measures and Total Fitness for Occupation Rating in Mechanic Learner Classes. (N = 80)**

Predictive Measure	Aircraft Instruments	Aircraft Engines	Aero Repair Mechanics	Machine Shop	Aircraft Electricity
Pressey Tests	.55 ± .05	.08 ± .08	.34 ± .06	.19 ± .09	.30 ± .09
Arithmetic Prognosis	.28 ± .07	.02 ± .08	.29 ± .07	.44 ± .08	.20 ± .10
Minnesota Paper Form Board	.18 ± .07	.18 ± .07	.03 ± .07	.20 ± .09	.22 ± .10
Finger Dexterity ...	−.02 ± .07	.10 ± .08	.14 ± .07	.17 ± .10	.31 ± .10
Tweezer Dexterity ..	.32 ± .06	.04 ± .08	.19 ± .07	.00 ± .11	.22 ± .10
Tridim. Structural Visualization	.26 ± .07	−.07 ± .08	.00 ± .07	.00 ± .10	.38 ± .09
Mechanical Comprehension ...	.35 ± .07	.11 ± .08	.30 ± .08	.35 ± .09	.41 ± .08
Placing Test	.07 ± .08	.11 ± .08	.15 ± .07	.01 ± .11	−.03 ± .11
Turning Test	.12 ± .08	.12 ± .08	.11 ± .07	.15 ± .10	.17 ± .10

* All tests were not given in all instances, thus making for a slight variation in the size of the sample.

Aircraft Instruments. Table 3 shows that some tests are fairly valid predictors of rated achievement in the Aircraft Instruments class. The *Pressey* tests, a measure of abstract intelligence, correlate .55 ± .05 with the Total Fitness rating. The *Mechanical Comprehension, Arithmetic Prognosis, Tridimensional Structural Visualization* and *Tweezer Dexterity* tests also produce fairly significant correlations. The correlations of the criterion with the other tests, *Revised Minnesota Paper Form Board, Finger Dexterity, Placing and Turning,*

indicate no usefulness in the prediction of Aircraft Instrument achievement during the learning period.

Aircraft Engines Overhaul. For the Aircraft Engines class, correlations between tests and ratings yield quite different results from the ones shown for the Instruments class. With the highest correlation being .18 ± .07 on the *Revised Minnesota Paper Form Board,* one can logically conclude that none of the tests significantly predict rated achievement in the Aircraft Engines class.

One factor which may have tended to lower the correlations in this group is that the amount of previous related experience was not a controlled factor. It is agreed by most industrialists that all classes involving the internal combustion engine are similar and related. Many individuals entering the Aircraft Engines class had been trained in Automobile Engine Mechanics prior to their enrollment. For this reason it is entirely possible that the previous training factor outweighed the aptitude factor, thus lowering the predictive significance of the tests in this situation. If actual job ratings were available, one might find that in due time the factor of previous line training would tend to equalize itself among all individuals, and aptitude might prove the distinguishing factor.

Aero Repair Mechanics. Results of correlations for the Aero Repair Mechanics class, presented in Table 3, show, with less significance, the same tendency that was exhibited in the Aircraft Instruments class. That is, tests involving abstract concepts again prove to be the best predictors of rated achievement. The *Pressey, Mechanical Comprehension,* and *Arithmetic Prognosis* tests bring correlations of .34 ± .06, .30 ± .07 and .29 ± .06, respectively. Although these correlations are low, they are slightly significant. With the exception of the *Tweezer Dexterity Test,* all manipulative tests and the *Paper Form Board Test* indicate that they could not be relied upon to predict achievement in this situation.

It is reasonable to expect these results in the Aero Repair training period, because there is much memory work and the situation is largely one involving verbal learning. No task becomes sufficiently routine in the primary training period for

one to impress an instructor with speed and dexterity on the job.

Machine Shop. Correlations of tests with Machine Shop rated achievement (Table 3) again show two of the tests involving problems of abstract thinking to be the best predictors. These are the *Arithmetic Prognosis* and *Mechanical Comprehension* tests. No other correlation coefficients show predictive significance of tests for machine work.

Aircraft Electricity. Four of the tests in Table 3 show indications of being rather significant as predictors of rated electrical achievement. Again, two of three abstract concepts tests prove their worth in the role of prediction. The *Pressey* and *Mechanical Comprehension* tests reveal correlations of $.30 \pm .09$ and $.41 \pm .08$. This shows that the *Mechanical Comprehension* test is a fairly significant predictor. The correlation of $.31 \pm .10$ for the *Finger Dexterity* test shows some relationship between the actual work processes and finger dexterity.

Multiple correlations. The tests estimated to be best for predicting achievement in each of the Mechanic Learner classes

TABLE 4

Multiple Correlations of Significant Predictive Devices for Each of the Mechanic Learner Classes

Class	Best combination of tests	Multiple correlation using best two tests	Multiple correlation using best predictive battery
Aircraft Instruments	Pressey Tests (.55) Mechanical Comprehension (.35) Tweezer Dexterity (.32) Arithmetic Prognosis (.28)	.61	.64
Aircraft Engines	No tests resulted in significant predictions		
Aero Repair Mechanics	Pressey Tests (.34) Mechanical Comprehension (.30) Arithmetic Prognosis (.29)	.42	.46
Machine Shop	Arithmetic Prognosis (.44) Mechanical Comprehension (.35)	.49	.49
Aircraft Electricity	Mechanical Comprehension (.41) Tri. Structural Visualization (.38) Finger Dexterity (.31) Pressey Tests (.30)	.54	.60

are indicated in Table 4. Here are shown correlation coefficients for single tests showing some predictive significance, multiple correlation coefficients for the best two tests, and multiple correlation coefficients for the best predictive battery of tests.

The best predictor for achievement in Aircraft Instruments was the *Pressey* tests, which brought a correlation of .55. Adding the *Mechanical Comprehension Test* resulted in a multiple coefficient of .61, and by use of all four tests, the correlation was raised to .64.

Since the single tests used for the Aircraft Engines Overhaul class showed no predictive significance, multiple correlations derived from them would also probably have no predictive value.

In the Aero Repair Mechanics class, a multiple correlation of .42 was obtained from the *Mechanical Comprehension* and *Pressey* tests. By the addition of a third test, the *Arithmetic Prognosis,* the multiple correlation was raised to .46.

The Machine Shop class had only two tests showing predictive significance. These tests, *Mechanical Comprehension* and *Arithmetic Prognosis,* yielded a multiple correlation coefficient of .49.

The best two predictors in the Aircraft Electricity class were *Mechanical Comprehension* and the *Tridimensional Structural Visualization* tests. They resulted in a multiple correlation of .54. Adding the *Finger Dexterity* and *Pressey* tests raised the multiple correlation to .60.

In general, it may be noted from Table 3 that a combination of the best two tests produces a validity nearly as high as the multiple correlations of three and four tests.

THE USE OF CIVIL SERVICE EXAMINATION BULLETINS AS A COUNSELING AID[1]

ARTHUR BURTON
California State Personnel Board

THE majority of civil service and public personnel agencies issue bulletins in one form or other as a means of recruiting personnel. Occupational counselors examine these bulletins to learn of positions available in the jurisdiction and then bring these job opportunities to the attention of clients whom they consider qualified for the position. This function of civil service bulletins as a means of job placement is well known. There is, however, an additional and equally important function of civil service examination bulletins which is not so well known and which this paper is intended to elaborate.

While civil service bulletins vary in form and content from one public personnel jurisdiction to another, they all contain essentially information designed to acquaint the reader with the details of the position and to induce him to apply for it. This information usually consists of the title of the position or class, description of the duties and responsibilities, the entrance salary and salary range, the entrance requirements, the examination date, final date for filing applications, and the location and conditions of employment. The nature of the qualifying examination—whether by written examination, performance test, oral interview, or a combination of these—and a statement of the knowledges, abilities, and personality requirements required for the position may also be given on the bulletin.

The job information which appears on a bulletin has more than job placement value to the counselor when it is considered how this information is derived originally. Information pro-

[1] I am indebted to Nova J. Beal for reviewing the manuscript.

vided on bulletins comes from class specifications which have been prepared by personnel technicians as a basic component of the jurisdiction's position-classification plan. Under this classification plan, there is a written specification for each position class and oftentimes for each grade of position within the class. The class specification contains, similar to the bulletin which is derived from it, a class title, job description[2] and definition, minimum education and experience requirements, personality requirements necessary for success on the job, and the knowledges, skills and abilities required. The salary range is obtained from the jurisdiction's salary evaluation plan. These components are determined by job surveys according to modern concepts and techniques of position classification. In the degree to which job and position analysis are utilized in a jurisdiction, the findings of salary determination, minimum entrance levels, requisite knowledges and abilities, and job descriptions are a valuable source of "job description" for the counselor, and this information is easily available and without cost.

As a source of occupational information in the form of "job descriptions," the civil service bulletin is in most instances comparable to "job descriptions" available from other sources. For professional and sub-professional positions where precise methods of job study such as time or motion study cannot strictly apply, they are often superior and in many cases are the only "job descriptions" available of acceptable merit. There is the further advantage that these bulletins represent "job descriptions" standardized and applicable to local conditions in the fact that such jobs actually exist in the geographical vicinity. Job descriptions, as for example those set up by the Occupational Analysis Section of the United States Employment Service, are standardized to represent a typical or average job on a nation-wide basis and as such suffer from some degree of

[2] The concept of "job description" is used in a general and specific sense in this paper. Job description, used specifically, implies a mere description of the duties and responsibilities of the job; whereas the general concept of "job description"—given in quotes—includes a complete picture of the job in all of its manifestations, and includes as a part of it the specific description of duties and responsibilities.

sampling error. In any given community, the actual job will vary from the "job description" to a greater or lesser degree.

In many public personnel jurisdictions it is agency policy, even where not mandatory by law, that journeymen such as plumbers, electricians, and carpenters be paid the prevailing wage rate in the state, county, city, or locality as the case may be. This prevailing rate is determined by the jurisdiction after survey of a sampling of all the positions in the area. The established prevailing rates are varied from time to time to conform to changes in wage rates in the community. A counselor by inspection of the prevailing rate stated on a civil service bulletin can ascertain at a glance the index of the wage paid for that class in industry. Where salaries are not prevailing rates, as for professional and clerical positions, the salary on the bulletin still represents an index of the wage that will attract an employee in the community. As a general index of wage rates in a community, other than prevailing rates, salaries of public personnel jurisdicitions are subject to variation from a true index and are inclined to be lower at certain levels than industrial and commercial rates for comparable positions.

In considering training for an occupation or in actual placement, a counselee desires to know not only what salary he may expect at the start but what the promotional possibilities are and what preparation is required to qualify for promotion. This information is often unavailable to the counselor, partly because promotion is contingent upon numerous factors in addition to opportunity and skill. Civil service bulletins collected over a period of years not only indicate entrance salary levels and skills but also indicate promotional levels and the skills required to attain them and the coordinate remuneration.

It is now apparent that the purpose of this note is to point out the value of collecting and utilizing in a consistent way the announcements of public personnel jurisdictions as a source of "job descriptions." If these bulletins are collected consistently and systematically, a procedure involving little expense or difficulty, the counselor could eventually have available an approximate replica of each specification contained in the files of

the public personnel jurisdiction. In the case of the California State Personnel Board, this would amount to approximately 1,600 individual class specifications. This file could then be used, as pointed out above, both as an aid in placement and in the same manner as other "job descriptions." When changes in job functions and duties occur, as in the lowering of entrance requirements and the "downgrading" of positions due to shortage of personnel, these changes will be reflected in contemporary bulletins and the extent of the change noted in comparing contemporary with past bulletins for the same class.

The validity of civil service bulletins as "job descriptions" will vary from jurisdiction to jurisdiction in accordance with the preciseness and methodology of job survey, which is essentially a basic element in the adequacy of the jurisdiction's position-classification plan. Counselors will have to determine empirically the relative validity of the bulletins as "job descriptions" and the degree to which they can be applied to their program.

MEASUREMENT ABSTRACTS*

Allen, Mary Elizabeth and Young, Florence M. "The Constancy of the Intelligence Quotient as Indicated by Re-Tests of 130 Children." *Journal of Applied Psychology,* XXVII (1943), 41–60.

This study was made to determine the constancy of the I.Q. and its relationship to sex, age, tests used for the testing and retesting, and the length of the interval between tests. It was found that (1) superior children (115 I.Q. or above) showed less change on retest than those below 115; (2) there was no significant I.Q. difference when the Stanford-Binet was used for both tests; (3) there was a significant difference when the *Merrill-Palmer* test was followed by the *Stanford-Binet;* (4) the group with the longest interval between test and retest (8 years, 10 months) showed the largest mean gain; (5) boys showed a greater change than girls; and (6) 2½ times as many subjects gained in I.Q. as lost. Fluctuation was often found to be a result of insecurity, illness, frequent changes of residence, conflict in the home, treatment prolonging immaturity, and frequent changes of schools. *Joan Murray.*

Bennett, George K. "Distribution of Scores from Revisions of Army Alpha." *Journal of Applied Psychology,* XXVII (1943), 150–151.

Hay and Blakemore report that the *Nebraska Revision of Alpha* given to 71 bank clerks produced a "strong negative skew." Wells' *Modified Alpha Form 9,* standardized upon some thousands of high-school students, indicates the absence of this characteristic. The statistics of Form 9 are so similar to those of Forms 8, 7, and 5 that these Forms are also free of this characteristic. A table is given presenting the percentile

* Edited by Forrest A. Kingsbury.

equivalents of high-school seniors drawn from 23 secondary schools in all areas of the United States to whom the *Modified Alpha Form 9* was administered. *M. H. Groves.*

Bennett, George K. and Gear, Richard A. "Mechanical Comprehension and Dexterity." *Personnel Journal,* XXII (1943), 12–17.

A mechanical ability battery, consisting of five mental and motor tests, was given to 40 machine operators in a recently converted ordnance-material factory. Employees were rated on a five-point scale by their shop supervisors. 91 per cent of the employees rated "excellent" scored high (above 60th percentile) on the tests, and 75 per cent of those rated "good" made similarly high scores. Applicants were also tested and those hired who made scores above the 60th percentile of the employees' scores. These new employees were later rated by their supervisors, only 8 per cent being rated "below average." The test also revealed individuals in low-level jobs who had potentialities for more difficult work. Further analysis showed that good selection could be done with only the *Mechanical Comprehension* and *Hand Tool Dexterity* tests, and the others were therefore dropped. *Joan Murray.*

Berdie, Ralph F. "A Procedure for Obtaining Representative Local Norms." *Journal of Consulting Psychology,* VII (1943), 160–163.

Local norms and the norms supplied by a test publisher may differ considerably. A procedure for obtaining representative local norms based on the intercorrelations between tests given to a group which is representative of the population on whom the test is to be used is described. Results show that local norms are more significant in interpreting scores being used for a particular purpose, such as chances for scholastic success in a specific college. *Lorraine Bouthilet.*

Blum, Milton L. "Selection of Sewing Machine Operators." *Journal of Applied Psychology,* XXVII (1943), 35–40.

Fifty sewing machine operators in a glove factory, 25 from

the highest quartile and 25 from the lowest quartile of earnings on a piece-rate basis, were tested on the *MacQuarrie Test for Mechanical Ability,* and on the *Blum Sewing Machine Test.* The latter consists of Test A, sewing on a zig-zagged line, and Test B, sewing between parallel zig-zagged lines; the score is the number of needle perforations away from the line and the time to complete the task. Criteria were production records, and the supervisors' ratings. The *MacQuarrie* test was found to be an unreliable indicator on the basis of the two criteria. The *Blum* tests show that the relation existing between the production records and the scores on the sewing machine test is positive. The correlation between time on the tests and the production record is about – .40. These results indicate that the *Blum* tests afford a valid measure to select good operators who are experienced. Whether it can be useful as an aptitude test is as yet undetermined. *M. H. Groves.*

Canady, Herman G. "The Problem of Equating the Environment of Negro-White Groups for Intelligence Testing in Comparative Studies." *Journal of Social Psychology,* XVII (1943), 3–13.

In spite of many attempts at solution, the problem of equating environments between Negro and white groups remains unsolved. The author reviews the various methods which have been used, pointing out the inadequacies of all of them. He concludes that it is doubtful if justifiable comparisons can be made unless there is a change in the social and economic conditions of the Negro. *K. S. Yum.*

Driver, R. S. "The Value of Psychological Tests in Industry." *Personnel,* XIX (1943), 656–664.

In response to the need for better methods of selection, induction, training, etc., of employees in industry a discussion of the use of tests for these purposes is given. The four methods described and evaluated are (1) use of tests which have been developed and standardized to meet the requirements of a particular situation, (2) the "clinical" method, (3) the "follow-up" method, (4) use of tests with known success in other situations. *Lorraine Bouthilet.*

Ferguson, Leonard W. "The Effect of a Second Administration of an Employment Test." *Journal of Applied Psychology,* XXVII (1943), 170–175.

Applicants for jobs with life insurance companies had the opportunity to take the same mental alertness test more than once. A comparison of scores on the two tests shows a mean increase of 12 points. Those who were hired showed a greater increase than those who were not hired. Those who made low scores on the first test but increased their scores enough on the second test to be hired seemed to be the less successful employees. It is suggested that the critical score by which acceptance for the job is made be revised an appropriate amount in cases where the applicant previously has taken the test. *Lorraine Bouthilet.*

Garrett, Henry E. "The Discriminant Function and Its Use in Psychology." *Psychometrika,* VIII (1943), 65–79.

R. A. Fisher's method of determining a set of weights for the linear combination of the scores in a battery of tests is explained and illustrated. The relationship of this method to combination by multiple regression methods is shown. (Courtesy *Psychometrika.*)

Ghiselli, Edwin E. "The Use of the Minnesota Rate of Manipulation and the O'Connor Finger Dexterity Tests in the Selection of Package Wrappers." *Journal of Applied Psychology,* XXVII (1943), 33–34.

Blum and Candee found the standard procedure of selecting package wrappers on the basis of the *Minnesota Rate of Manipulation* and the *O'Connor Finger Dexterity Tests* an unreliable one, with a validity coefficient for the former of about .20, and for the latter of about .00. Investigations of the U. S. Employment Service corroborated these findings on the Minnesota tests. The author checked these findings by correlating each of the two tests with ratings on 42 seasonal wrappers at a department store, all of whom had taken both tests. The validity coefficients of the placing and turning tests were, respectively, – .10 and – .02, while that of the finger dexterity test

was + .02. While these tests are useless for such selection, the name and number comparison tests have been found useful and tapping and dotting tests, and paper and pencil spatial relations tests may be helpful. *M. H. Groves.*

Goodman, Charles H. "Factorial Analysis of Thurstone's Seven Primary Abilities." *Psychometrika,* VIII (1943), 121–129.

Intercorrelation coefficients among Thurstone's seven primary mental abilities scores were obtained from scores of 170 freshmen engineering students on the experimental edition of the *Primary Mental Abilities Tests.* These correlation coefficients were factored to four factors, interpreted as a general factor, a reasoning factor, a verbal factor, and, tentatively, a specific memory factor. The finding of a general factor for a college population corroborates the Thurstones' finding of a general factor for eighth-grade children. (Courtesy *Psychometrika.*)

Greene, Edward B. "An Analysis of Random and Systematic Changes with Practice." *Psychometrika,* VIII (1943), 37–52.

Six motor tests and six nonverbal tests were administered four times to the same subjects. Subjective reports of the subjects are discussed, changes in mean scores and in variability and score correlations from trial to trial are surveyed, and factor analyses of results on the first and fourth trials are presented and compared. Implications of the findings with respect to correction for attenuation are pointed out. (Courtesy *Psychometrika.*)

Hebb, D. O. and Morton, N. W. "The McGill Adult Comprehension Examination: 'Verbal Situation' and 'Picture Anomaly' Series." *Journal of Educational Psychology,* XXXIV (1943), 16–25.

A verbal and non-verbal intelligence test for adults were designed to supplement the more abstract problems of other tests. In the "verbal situation" test the subject is required to

give a common-sense judgment of an ordinary situation; in the "picture anomaly" test he is asked to point out the incongruities or ludicrousness in a picture. Reliability coefficients above .90 are reported, as well as relatively good agreement with other estimates of intelligence. It appears that the tests measure "a kind of cultural comprehension which may be important in itself, and that the degree of comprehension may provide a better basis for estimating or rating intelligence than many current tests." *Lorraine Bouthilet.*

Heston, Joseph E. "A Factor Analysis of Some Clinical Performance Tests." *Journal of Applied Psychology,* XXVII (1943), 135–149.

Following a summary of several studies done on performance tests, the author describes his own study of the primary factors in such tests. Subjects were 113 white male college students in their first quarter. From eighteen variables tested, three factors were drawn as significant in the light of the criteria. Two, spatial and speed, were readily seen and describable. The third was tentative and not adequately determined, although thought to be ability to grasp non-spatial Gestalt relationships. It was concluded that the battery used required an ability within the "unique" factor space rather than the common factor space; that only 33 per cent of the variance was accounted for by the three factors, the rest being explainable on the basis of unique and unrelated components; and that no common factor ("g") was found to pervade all the tests. *Joan Murray.*

Jurgensen, Clifford E. "Extension of the Minnesota Rate of Manipulation Test." *Journal of Applied Psychology,* XXVII (1943), 164–169.

Test procedure in the administration of the *Minnesota Rate of Manipulation Test* was extended to give seven speed measures in addition to those ordinarily used. The additional procedures yielded higher reliability and validity coefficients and lower intercorrelations between tests. It is recommended that employers using the test not limit themselves to the customary ways of presenting and scoring it. *Lorraine Bouthilet.*

Reyburn, H. A. and Taylor, J. G. "On the Interpretation of Common Factors: A Criticism and a Statement." *Psychometrika,* VIII (1943), 53–64.

The concept of simple structure is criticized for lack of objectivity and for failure to produce invariance of (a) factor loadings under change of battery, and (b) individual scores on primary traits under change of tests. It is also criticized on the grounds that simple structure yields at best the factors which were put in and that, by suitable manipulation of tests, any set of factors may be represented in a battery by tests which will yield a simple structure. A procedure for rotation is developed which locates the first rotated axis to pass through a cluster of tests which, by hypothesis, contain a common factor and to project all tests into a hyperplane orthogonal to this factor. The second factor is then located to pass through the projections of a second cluster of tests defining the second factor, and so on, until all hypothecated factors have been located. Any residual interrelations may then be rotated graphically to the most plausible arrangement. (Courtesy *Psychometrika.*)

Reyburn, H. A. and Taylor, J. G. "Some Factors of Temperament: A Re-Examination." *Psychometrika,* VIII (1943), 91–104.

By application of the technique presented earlier,* the factors found by the Guilfords in two studies of personality factors are reanalyzed. The rotation technique for the isolation of meaningful factors relates the factors of the two studies to each other through three common items, and to other factors and traits appearing in the literature of personality measurement. For each of four of the factors resulting from the analysis of the first battery there is close agreement with its counterpart from the analysis of the second battery with respect to loadings of the three common items. These factors are interpreted as perseverance, surgency, flexibility, and tension. Other factors appearing in only one or the other of the batteries are tentatively identified. (Courtesy *Psychometrika.*)

---

* *Psychometrika,* VIII (1943), 53–64.

Slater, Patrick. "The Use of Percentiles." *Occupational Psychology,* XVII (1943), 73–81.

This is a survey of the use of percentiles as a statistical tool in psychology. Methods of computation, interpretation, and comparison with the I.Q. are given. *K. S. Yum.*

Super, Donald E. and Haddad, C. "The Effect of Familiarity With an Occupational Field on a Recognition Test of Vocational Interest." *Journal of Educational Psychology,* XXXIV (1943), 103–109.

This article describes a study made to determine whether familiarity with an occupational field influences scores on a recognition test of vocational interest. One hundred and thirty-five high-school boys were used as subjects, and four tests were used, the *Super-Older Test of Vocational Interest,* the *Michigan Vocabulary Profile Test* (modified), the *Strong Vocational Interest Blank,* and the *Otis Q.S. Mental Ability Test.* The degree of familiarity with an occupational field had no effect upon memory scores for material dealing with that occupational field. Since it appears that intelligence plays a negligible part in the retention of material such as this, the author concludes that the memory scores are probably a reflection of the focusing of attention upon objects or activities which challenge the individual's interest. *Joan Murray.*

Tiffin, Joseph and Lawshe, C. H., Jr. "The Adaptability Test: A Fifteen-Minute Mental Alertness Test for Use in Personnel Administration." *Journal of Applied Psychology,* XXVII (1943), 152–163.

This fifteen-minute test was designed to predict mental alertness and adaptability primarily in the clerical field. Two forms, each of thirty-five items, were compared and their reliability, determined by both the split-half and test-retest methods, was found to approach .90. The scores of high-rated and low-rated clerical workers were found significantly different. Those persons who made high scores were well adapted to clerical work, and those who made low scores were on the whole well adapted to work of a routine nature. Individuals

rated as competent inspectors or other routine workers made lower scores on the test than those who were rated as poor inspectors and routine workers. The authors also present specific recommendations for placement on the basis of specific scores. *Joan Murray.*

Wittenborn, John Richard. "Factorial Equations for Tests of Attention." *Psychometrika,* VIII (1943), 19–35.

Tests designed to measure what was conceived to be attention were found by factor analysis to involve a factor which is independent of the factors of rote memory, visual-space, number, or perception. To a large extent, at least, this attention factor is independent of content and of mode of presentation of test material. The tests in which the variance is mainly dependent upon this factor are those involving a high degree of sustained or relatively continuous mental effort. (Courtesy *Psychometrika.*)

NEWS NOTES*

Mrs. Mary B. Bondurant, Director, Division of Placement, University of Georgia, has been granted a two-year leave of absence beginning July 1, 1943. This was necessitated by the fact that her husband has been on active duty with the Army Air Corps for some time and Mrs. Bondurant has taken over his business. She has also resigned as Secretary of the National Institutional Teacher Placement Association.

Dr. Blanche Carrier, previously Dean of Women and Professor of Psychology at Arizona State Teachers College, Tempe, Arizona, has resigned to become Director of Week-Day Schools and Parent Education for the Northern California Council of Religious Education, 83 McAllister Street, San Francisco.

Lieut. Colonel Lysle W. Croft, A.G.D., is now stationed at Headquarters, Replacement and School Command, Army Ground Forces, Birmingham, Alabama.

Margaret D. Gordy, Placement Director, Wilson College, Chambersburg, Pa., is spending the summer serving as counselor and liaison officer for the apprentice training group of the Grumman Aircraft Engineering Corporation. The group will study at Columbia University School of Engineering from June 20 to August 20 as part of the program of professional and technical training leading to positions in the engineering department for women who are college graduates.

Anna M. Hanson, formerly Teaching Appointments Secretary at Yale Graduate School, has accepted the position of Director of Placement at Simmons College, Boston, Massachusetts.

Lt. Comdr. O. C. Held, U.S.N.R., Colgate University, Hamilton, N. Y. After two years of service with the Neuropsychiatric Unit, Lt. Comdr. Held has been transferred to administrative work in the college training program at Colgate University.

E. R. Henry is on leave from his duties as Chairman, Department of Psychology, New York University, and is now located with the Personnel Research Section, Classification and Replacement Branch, Adjutant General's Office, 270 Madison Avenue, New York City.

* News items concerning members of the American College Personnel Association should be sent to Grace E. Manson, Northwestern University, Evanston, Illinois.

Miss Florence Jackson, Vocational Consultant, is planning a new series of talks on the topic "Toward New Horizons" to be used in advising preparatory school and college students.

W. C. Johnson, Director of Vocational Guidance and Placement at Virginia Polytechnic Institute, has recently taken on the additional duties of Acting Director of Admissions.

Lieut. (j.g.) Howard E. Miller, U.S.N.R., entered service in the Navy last August. He was previously Registrar and Director of Personnel, Illinois Wesleyan University, Bloomington, Illinois.

Mrs. Marjorie Nield, Director, Vocational Office, Smith College, has recently published a bulletin useful to Smith students in choosing and preparing for war work entitled "College Women in War Work." It gives the nature of the work, the training required, additional training desired, and a job description.

Dr. Charles L. Odom, Associate Professor of Psychology, Southwestern Louisiana Institute, has been a visiting professor at the George Peabody College for Teachers during the summer quarter.

Edwin F. Peters, formerly at William Woods College, Fulton, Missouri, is now in the United States Navy and attended the training course at Columbia University for educational administrative work in connection with the navy program in the colleges.

Warren T. Powell, Director Student Counseling, Boston University, Boston, Mass., has sent us an announcement of the two-week Institute on Guidance and Personnel with war-time guidance as its theme, which the Boston University School of Education is sponsoring.

Virginia J. Reinecke, who has been Assistant Counselor of Women at the University of Oklahoma, has been appointed Counselor of Women. Since the naval personnel in Norman numbers 15,000, the University has assumed many new responsibilities.

Lieut. (j.g.) Gwendolen G. Schneidler, U.S.N.R. (W.R.), on leave from the Testing Bureau at the University of Minnesota, is now stationed at the Bureau of Naval Personnel, Training Division, Standards and Curriculum Section, Washington, D. C., working on problems of selection and classification of enlisted personnel in the Navy and Waves.

H. E. Stone, Secretary, University Placement Service, West Virginia University, has an article entitled "Placement, Past, Present and Future" in the May issue of "School and College Placement."

Lieut. (j.g.) Reginald L. Swann, H–V(s), U.S.N.R., U. S. Naval Aviation Cadet Selection Board, Allen Building, Dallas, Texas, is one of two psychologists with this Board. His work consists of testing and interviewing aviation applicants. Mr. Swann was previously Professor of Psychology and Director of Student Personnel at Green Mountain Junior College, Poultney, Vermont.

Mrs. Chase Going Woodhouse, Institute of Women's Professional Relations, New London, Conn., is serving as Public Panel Member

for the New England Regional War Labor Board. For over a year she has also served as Consultant, National Roster of Scientific and Specialized Personnel, War Manpower Commission. In this capacity she has organized War Work Information Centers along the Atlantic coast and in St. Louis and Minneapolis, with others under way in the Middle West. These centers are sponsored by various women's organizations, are staffed by volunteers and work in very close cooperation with the USES and The War Manpower Area Director. They are directed primarily toward the recruitment of women with specialized training, but also serve as information centers for women who are desirous of making a contribution to the war but who have not been in the labor market and who are not certain as to how to proceed.

RECOMMENDATIONS OF THE COUNCIL OF GUIDANCE AND PERSONNEL ASSOCIATIONS[1]

THE Council of Guidance and Personnel Associations, representing 5,000 counselors of American youth in high schools and colleges, met January 15–16, 1943, with representatives of the War Manpower Commission, Armed Services, United States Office of Education, Bureau of Labor Statistics, American Council on Education, and representatives of labor and industry to consider the specialized responsibilities of counselors, deans, and other personnel workers in bringing about an early victory, a just and lasting peace, harmonious and productive group relations in our society, and the happiness and welfare of individuals. These ends are accomplished most effectively through the development and utilization of human resources. Counselors and personnel specialists, knowing something of the aspirations of individuals and their conscious and unconscious needs and motivations, exercise a directive influence, individual by individual, on the development and utilization of human resources.

More specifically, the task of personnel workers is

1. To help every young person make his maximum contribution to an early victory by attaining the best health and physical fitness possible for him; by developing good work habits and self-discipline, by planning a healthful and educational twenty-four hour schedule, and otherwise fulfilling his potentialities as far as possible; by choosing courses in physics, mathematics, pre-flight aeronautics, and other specialized subjects on the basis of individual ability to make a contribution in the field as well as on the basis of their value in preparing

[1] The detailed report of the conference may be obtained from R. F. Moore, 528 West 121st Street, New York City, or in the March, 1943, issue of the *Journal of the National Association of Deans of Women*, or in the April, 1943, issue of *Occupations*.

young men for induction; by undertaking an accelerated program only if his ability, previous preparation, health, and present outside contribution to the war effort warrant such concentrated effort, recognizing that indiscriminate choice or assignment will result in maladjustments rather than in well-trained workers; by understanding and keeping in mind the present and the future over-all occupational picture, which includes essential occupations, industry and agriculture, and community services such as teaching, nursing, social services, and food production, as well as new military, industrial, and agricultural demands; by being open-minded and flexible with respect to military requirements, pursuing carefully planned educational programs until new ones required by the emergency are evolved; by accepting and understanding his classification and assignment in the Armed Forces; by filling out the Educational Experience Summary Card recommended by the United States Office of Education to be taken by each youth to his induction or reception center; by preparing himself through a realistic and accurate self-appraisal to give information needed in the interviews held at the reception center.

2. To help every young person to do his part in effecting a just and lasting peace by maintaining, under war conditions, the highest possible standards of personal and social behavior, while recognizing the impact of the world crisis on his personal life; by understanding the aims, backgrounds, and issues of the war and the responsibilities of military service in any part of the world; by focusing attention on the potential contribution of each nation to a world federation, thus eliminating destructive hatred for any peoples; by making long-range plans and considering the social value as well as personal satisfaction in choosing his life work. The Council will continue to study and suggest methods of attacking postwar guidance problems anticipated in the United States and in other countries.

3. To help each young person to work cooperatively and constructively in groups, finding personal satisfaction in the success of the group; by practicing democratic procedures in class, club, and other school and college groups; by becoming more proficient in the discussion of current issues and social

problems within his scope of experience and comprehension, recognizing the unique contribution which each member of the group can make; by developing leaders who exemplify democratic procedures and who are not swayed by prejudices.

4. To help to create conditions which make effective guidance possible through:

(a) A curriculum which will meet the needs of students and of society in these changing years. Among the features to be included are those outlined in the High School Victory Corps; opportunities for exceptionally well-qualified youth to receive accelerated education for occupations requiring extended preparation; provision of scholarship funds if necessary; and part-time educational programs in cooperation with industrial organizations in order to help meet labor shortages and to provide selected students with educative work experiences while they are continuing their education. Even in areas of critical labor shortages young people should not be deprived of the education they need in order to develop their potential capacities, until all other sources of labor supply, such as minority groups, the handicapped, and women without small children to care for, have been utilized.

(b) A strong developmental guidance program in the elementary school.

(c) Personally qualified and well-trained personnel workers to serve as consultants in high schools and colleges, in rural counties, and in state departments. Adequate personnel services at other points where such services are needed, e.g., employment offices, social agencies, and industry.

(d) Desirable labor standards.

(e) Regulation and investigation of private vocational and correspondence schools.

(f) Better preparation, in teacher-education institutions, of teachers, counselors, and deans in the personnel point of view and guidance procedures.

5. To help survey and co-ordinate the services of community agencies—the schools, Public Employment Service, group work and case work agencies—by understanding and co-operating with these agencies, by bringing the Council recom-

mendations to their attention, by promoting the full use of community resources in providing constructive leisure-time activities for young people, including service men, by helping to select carefully and prepare girls to serve at centers for the entertainment of service men.

Through committees and further conferences the Council will endeavor to keep up-to-date with new developments and, as individuals and constituent groups, implement the recommendations made in the 1943 meeting.

By the Committee on Publications:
KARL M. COWDERY
FRANCES MAULE
RUTH STRANG, Chairman

e of
ac-
g to
the

ncil
and,
om-